PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

	INTERNATIONAL APPLICATION FORLIST	TLD V	J. 11 J. 11 11 11 11 11 11 11 11 11 11 11 11 11	
ļ	(51) International Patent Classification ⁶ :		(11) International Publication Number:	WO 99/47540
	C07H 21/04, C07K 1/00, 16/00, C12N 15/00, 1/20, C12P 21/06, A61K 38/00, C12Q 1/68, G01N 33/53, 33/566	A1	(43) International Publication Date: 2	23 September 1999 (23.09.99)
	(21) International Application Number: PCT/US (22) International Filing Date: 18 March 1999 (MD 20853 (US). ROSEN, Rolling Hill Road, Laytonsv Guo-Liang [CN/US]: 242 C	Craig, A. [US/US]; 22400 ville, MD 20882 (US). YU,

(30) Priority Data:		
60/078,566	19 March 1998 (19.03.98)	US
60/078,576	19 March 1998 (19.03.98)	US
60/078,573	19 March 1998 (19.03.98)	US
60/078,574	19 March 1998 (19.03.98)	US
60/078,579	19 March 1998 (19.03.98)	US
60/078,578	19 March 1998 (19.03.98)	US
60/078,581	19 March 1998 (19.03.98)	US
60/078,577	19 March 1998 (19.03.98)	US
60/078,563	19 March 1998 (19.03.98)	US
60/080,314	1 April 1998 (01.04.98)	US
60/080,312	1 April 1998 (01.04.98)	US
60/080,312	1 April 1998 (01.04.98)	US

- (71) Applicant (for all designated States except US): HUMAN GENOME SCIENCES, INC. [US/US]; 9410 Key West Avenue, Rockville, MD 20850 (US).
- (72) Inventors; and
 (75) Inventors/Applicants (for US only): RUBEN, Steven, M. [US/US]; 18528 Heritage Hills Drive, Olney, MD 20832

(US). NI, Jian [CN/US]; 5502 Manorfield Road, Rockville, MD 20853 (US). ROSEN, Craig, A. [US/US]; 22400 Rolling Hill Road, Laytonsville, MD 20882 (US). YU, Guo-Liang [CN/US]; 242 Gravett Drive, Berkeley, CA 94705 (US). YOUNG, Paul, E. [US/US]; 122 Beckwith Street, Gaithersburg, MD 20878 (US). FENG, Ping [CN/US]; 4 Relda Court, Gaithersburg, MD 20878 (US). SOPPET, Daniel, R. [US/US]; 15050 Stillfield Place, Centreville, MD 22020 (US). WEI, Ying-Fei [CN/US]; 242 Gravatt Drive, Berkeley, CA 94705 (US). ENDRESS, Gregory, A. [US/US]; 9729 Clagett Farm Drive, Potomac, MD 20854 (US). DUAN, Roxanne, D. [US/US]; 5515 Northfield Road, Bethesda, MD 20817 (US). KYAW, Hla [MM/US]; 520 Sugarbush Circle, Frederick, MD 21703 (US). EBNER, Reinhard [DE/US]; 9906 Shelburne Terrace #316, Gaithersburg, MD 20878 (US). LAFLEUR, David, W. [US/US]; 3142 Quesada Street, N.W., Washington, DC 20015 (US). OLSEN, Henrik, S. [DK/US]; 182 Kendrick Place #24, Gaithersburg, MD 20878 (US). SHI, Yanggu [CN/US]; 437 West Side Drive #102, Gaithersburg, MD 20878 (US). Gaithersburg, MD 20878 (US). Gaithersburg, MD 20878 (US). Gaithersburg, MD 20878 (US). SHI, Yanggu [CN/US]; 437 West Side Drive #102, Gaithersburg, MD 20878 (US).

- (74) Agents: BROOKES, A., Anders et al.; Human Genome Sciences, Inc., 9410 Key West Avenue, Rockville, MD 20850 (US).
- (81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TI, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TI, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

With international search report.

(54) Title: 95 HUMAN SECRETED PROTEINS

(57) Abstract

The present invention relates to novel human secreted proteins and isolated nucleic acids containing the coding regions of the genes encoding such proteins. Also provided are vectors, host cells, antibodies, and recombinant methods for producing human secreted proteins. The invention further relates to diagnostic and therapeutic methods useful for diagnosing and treating disorders related to these novel human secreted proteins.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE.	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Paso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	īL	Israel	MR	Mauritania	υG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	zw	Zimbabwe
CI	Côre d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon	***	Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
cz	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	u	Liechtenstein	SID	Sudan		
DK DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		
r.K.	ERMIN	Lan			- 0.		

5

10

15

20

25

30

1

95 Human Secreted Proteins

Field of the Invention

This invention relates to newly identified polynucleotides and the polypeptides encoded by these polynucleotides, uses of such polynucleotides and polypeptides, and their production.

Background of the Invention

Unlike bacterium, which exist as a single compartment surrounded by a membrane, human cells and other eucaryotes are subdivided by membranes into many functionally distinct compartments. Each membrane-bounded compartment, or organelle, contains different proteins essential for the function of the organelle. The cell uses "sorting signals," which are amino acid motifs located within the protein, to target proteins to particular cellular organelles.

One type of sorting signal, called a signal sequence, a signal peptide, or a leader sequence, directs a class of proteins to an organelle called the endoplasmic reticulum (ER). The ER separates the membrane-bounded proteins from all other types of proteins. Once localized to the ER, both groups of proteins can be further directed to another organelle called the Golgi apparatus. Here, the Golgi distributes the proteins to vesicles, including secretory vesicles, the cell membrane, lysosomes, and the other organelles.

Proteins targeted to the ER by a signal sequence can be released into the extracellular space as a secreted protein. For example, vesicles containing secreted proteins can fuse with the cell membrane and release their contents into the extracellular space - a process called exocytosis. Exocytosis can occur constitutively or after receipt of a triggering signal. In the latter case, the proteins are stored in secretory vesicles (or secretory granules) until exocytosis is triggered. Similarly, proteins residing on the cell membrane can also be secreted into the extracellular space by proteolytic cleavage of a "linker" holding the protein to the membrane.

Despite the great progress made in recent years, only a small number of genes encoding human secreted proteins have been identified. These secreted proteins include the commercially valuable human insulin, interferon, Factor VIII, human

growth hormone, tissue plasminogen activator, and erythropoeitin. Thus, in light of the pervasive role of secreted proteins in human physiology, a need exists for identifying and characterizing novel human secreted proteins and the genes that encode them. This knowledge will allow one to detect, to treat, and to prevent medical disorders by using secreted proteins or the genes that encode them.

Summary of the Invention

The present invention relates to novel polynucleotides and the encoded polypeptides. Moreover, the present invention relates to vectors, host cells, antibodies, and recombinant methods for producing the polypeptides and polynucleotides. Also provided are diagnostic methods for detecting disorders related to the polypeptides, and therapeutic methods for treating such disorders. The invention further relates to screening methods for identifying binding partners of the polypeptides.

15

20

25

30

10

5

Detailed Description

Definitions

The following definitions are provided to facilitate understanding of certain terms used throughout this specification.

In the present invention, "isolated" refers to material removed from its original environment (e.g., the natural environment if it is naturally occurring), and thus is altered "by the hand of man" from its natural state. For example, an isolated polynucleotide could be part of a vector or a composition of matter, or could be contained within a cell, and still be "isolated" because that vector, composition of matter, or particular cell is not the original environment of the polynucleotide.

In the present invention, a "secreted" protein refers to those proteins capable of being directed to the ER, secretory vesicles, or the extracellular space as a result of a signal sequence, as well as those proteins released into the extracellular space without necessarily containing a signal sequence. If the secreted protein is released into the extracellular space, the secreted protein can undergo extracellular processing

5

10

15

20

25

30

to produce a "mature" protein. Release into the extracellular space can occur by many mechanisms, including exocytosis and proteolytic cleavage.

In specific embodiments, the polynucleotides of the invention are less than 300 kb, 200 kb, 100 kb, 50 kb, 15 kb, 10 kb, or 7.5 kb in length. In a further embodiment, polynucleotides of the invention comprise at least 15 contiguous nucleotides of the coding sequence, but do not comprise all or a portion of any intron. In another embodiment, the nucleic acid comprising the coding sequence does not contain coding sequences of a genomic flanking gene (i.e., 5' or 3' to the gene in the genome).

As used herein, a "polynucleotide" refers to a molecule having a nucleic acid sequence contained in SEQ ID NO:X or the cDNA contained within the clone deposited with the ATCC. For example, the polynucleotide can contain the nucleotide sequence of the full length cDNA sequence, including the 5' and 3' untranslated sequences, the coding region, with or without the signal sequence, the secreted protein coding region, as well as fragments, epitopes, domains, and variants of the nucleic acid sequence. Moreover, as used herein, a "polypeptide" refers to a molecule having the translated amino acid sequence generated from the polynucleotide as broadly defined.

In the present invention, the full length sequence identified as SEQ ID NO:X was often generated by overlapping sequences contained in multiple clones (contig analysis). A representative clone containing all or most of the sequence for SEQ ID NO:X was deposited with the American Type Culture Collection ("ATCC"). As shown in Table 1, each clone is identified by a cDNA Clone ID (Identifier) and the ATCC Deposit Number. The ATCC is located at 10801 University Boulevard, Manassas, Virginia 20110-2209, USA. The ATCC deposit was made pursuant to the terms of the Budapest Treaty on the international recognition of the deposit of microorganisms for purposes of patent procedure.

A "polynucleotide" of the present invention also includes those polynucleotides capable of hybridizing, under stringent hybridization conditions, to sequences contained in SEQ ID NO:X, the complement thereof, or the cDNA within the clone deposited with the ATCC. "Stringent hybridization conditions" refers to an

4

overnight incubation at 42° C in a solution comprising 50% formamide, 5x SSC (750 mM NaCl, 75 mM sodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 μ g/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65°C.

5

10

15

20

25

30

Also contemplated are nucleic acid molecules that hybridize to the polynucleotides of the present invention at lower stringency hybridization conditions. Changes in the stringency of hybridization and signal detection are primarily accomplished through the manipulation of formamide concentration (lower percentages of formamide result in lowered stringency); salt conditions, or temperature. For example, lower stringency conditions include an overnight incubation at 37°C in a solution comprising 6X SSPE (20X SSPE = 3M NaCl; 0.2M NaH₂PO₄; 0.02M EDTA, pH 7.4), 0.5% SDS, 30% formamide, 100 ug/ml salmon sperm blocking DNA; followed by washes at 50°C with 1XSSPE, 0.1% SDS. In addition, to achieve even lower stringency, washes performed following stringent hybridization can be done at higher salt concentrations (e.g. 5X SSC).

Note that variations in the above conditions may be accomplished through the inclusion and/or substitution of alternate blocking reagents used to suppress background in hybridization experiments. Typical blocking reagents include Denhardt's reagent, BLOTTO, heparin, denatured salmon sperm DNA, and commercially available proprietary formulations. The inclusion of specific blocking reagents may require modification of the hybridization conditions described above, due to problems with compatibility.

Of course, a polynucleotide which hybridizes only to polyA+ sequences (such as any 3' terminal polyA+ tract of a cDNA shown in the sequence listing), or to a complementary stretch of T (or U) residues, would not be included in the definition of "polynucleotide," since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone).

The polynucleotide of the present invention can be composed of any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. For example, polynucleotides can be composed of

5

single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, the polynucleotide can be composed of triple-stranded regions comprising RNA or DNA or both RNA and DNA. A polynucleotide may also contain one or more modified bases or DNA or RNA backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications can be made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically, or metabolically modified forms.

5

10

15

20

25

30

The polypeptide of the present invention can be composed of amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres, and may contain amino acids other than the 20 gene-encoded amino acids. The polypeptides may be modified by either natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched, for example, as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched, and branched cyclic polypeptides may result from posttranslation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine, formation of pyroglutamate, formylation,

gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. (See, for instance, PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993); POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York, pgs. 1-12 (1983); Seifter et al., Meth Enzymol 182:626-646 (1990); Rattan et al., Ann NY Acad Sci 663:48-62 (1992).)

"SEQ ID NO:X" refers to a polynucleotide sequence while "SEQ ID NO:Y" refers to a polypeptide sequence, both sequences identified by an integer specified in Table 1.

"A polypeptide having biological activity" refers to polypeptides exhibiting activity similar, but not necessarily identical to, an activity of a polypeptide of the present invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. In the case where dose dependency does exist, it need not be identical to that of the polypeptide, but rather substantially similar to the dose-dependence in a given activity as compared to the polypeptide of the present invention (i.e., the candidate polypeptide will exhibit greater activity or not more than about 25-fold less and, preferably, not more than about tenfold less activity, and most preferably, not more than about three-fold less activity relative to the polypeptide of the present invention.)

Polynucleotides and Polypeptides of the Invention

25

30

5

10

15

20

FEATURES OF PROTEIN ENCODED BY GENE NO: 1

This gene is expressed primarily in anergic T cells and merkel cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune disorders and inflammatory diseases. Similarly, polypeptides

7

and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

5

15

20

25

30

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 108 as residues: Ala-55 to Gln-64.

The tissue distribution in T-cells and merkel cells indicates that the protein products of this gene are useful for the diagnosis and/or treatment of immune system diseases. Furthermore,

Expression of this gene product in T-cells indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:11 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2329 of SEQ ID NO:11, b is an integer of 15 to 2343, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:11, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 2

10

15

20

In specific embodiments, polypeptides of the invention comprise the following amino acid sequences: IPENRRPASXCTWSMWTSRTTTRRPPWGRFSSVSSASV SSTRKTWRTRSTSCCRSSRRRVAAPFCTPSASTEPSARMEPPLELPVVHTFSFL TFVFTYRCSAGDGSITQINCAYEMGEEMPKRQMKAIKFLLFHFYL (SEQ ID NO:205), IPENRRPASXCTWSMWTSRTTTRRPPWGRFSSVSSASVSST (SEQ ID NO:206), RKTWRTRSTSCCRSSRRRVAAPFCTPSASTEPSARMEPPLELP (SEQ ID NO:207), and/or VVHTFSFLTFVFTYRCSAGDGSITQINCAYEMGEEMPKRQ MKAIKFLLFHFYL (SEQ ID NO:208). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in placental, brain and breast tissues, and to a lesser extent in T cells and tumors.

Therefore, polynucleotides and polypeptides of the invention are useful as

reagents for differential identification of the tissue(s) or cell type(s) present in a

biological sample and for diagnosis of diseases and conditions which include, but are
not limited to, neurodegenerative and/or endocrine disorders and neoplasias, or
developmental disorders. Similarly, polypeptides and antibodies directed to these
polypeptides are useful in providing immunological probes for differential

identification of the tissue(s) or cell type(s). For a number of disorders of the above
tissues or cells, particularly of the neurodegenerative, developing, endocrine and

9

immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., brain, endocrine, immune, developing, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

5

10

15

20

25

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 109 as residues: Ala-55 to Asn-60, Lys-65 to Met-71, Leu-75 to Asn-86, Asp-93 to Asp-110, Leu-130 to Cys-138, Gln-149 to Glu-154, Thr-172 to Ile-179, Glu-185 to Arg-192.

The tissue distribution in breast and brain tissues indicates that the protein products of this gene are useful for the diagnosis and/or treatment of endocrine disorders, neurodegenerative disorders, developmental disorders, immune system diseases and neoplasias. The tissue distribution in placental tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of disorders of the placenta. Specific expression within the placenta indicates that this gene product may play a role in the proper establishment and maintenance of placental function. Alternately, this gene product may be produced by the placenta and then transported to the embryo, where it may play a crucial role in the development and/or survival of the developing embryo or fetus.

Expression of this gene product in a vascular-rich tissue such as the placenta also indicates that this gene product may be produced more generally in endothelial cells or within the circulation. In such instances, it may play more generalized roles in vascular function, such as in angiogenesis. It may also be produced in the vasculature and have effects on other cells within the circulation, such as hematopoietic cells. It may serve to promote the proliferation, survival, activation, and/or differentiation of hematopoietic cells, as well as other cells throughout the body. Likewise,

Expression of this gene product in T-cells indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be

involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses).

5

10

15

20

25

30

Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Alternatively, the tissue distribution in brain tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, or sexually-linked disorders. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:12 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1163 of SEQ ID NO:12, b is an

11

integer of 15 to 1177, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:12, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 3

The translation product of this gene shares sequence homology with bovine beta-mannosidase, which is thought to be important in lysosomal catabolism of glycoproteins. See, for example, J. Biol. Chem. 270, 3841-3848 (1995), incorporated herein by reference in its entirety. Based on the sequence similarity between these proteins the translation product of this gene will sometimes hereinafter be reffered to as human beta-mannosidase. Human beta-mannosidase is expected to share certain biological activities, particularly enzymatic activities, with bovine beta-mannosidase. Such activities may be assayed by methods known in the art, described in J. Biol. Chem. 270, 3841-3848 (1995), and/or disclosed elsewhere herein.

In specific embodiments, polypeptides of the invention comprise the following amino acid sequences: HPSIIIWSGNNENEEALMMNWYHISFTDRPIYIKDYVTL YVKNIRELVLAGDKSRPFITSSPTNGAETVAEAWVSQNPNSNYFGDVHFYDYI SDCWNWKVFPKARFASEYGYQSWPSFSTLEKVSSTEDWSFNSKFSLHRQHH EGGNKQMLYQAGLHFKLPQSTDPLRTFKDTIYLTQVMQAQCVKTETEFYRRS RSEIVDQQGHTMGALYWQLNDIWQAPSW (SEQ ID NO:209), and/or

20 VRVHTWS

5

10

15

25

SLEPVCSRVTERFVMKGGEAVCLYEEPVSELLRRCGNCTRESCVVSFYLSAD HELLSPTNYHFLSSPKEAVGLCKAQITAIISQQGDIFVFDLETSAVAPFVWLDV GSIPGRFSDNGFLMTEKTRTILFYPWEPTSKNELEQSFHVTSLTDIY (SEQ ID NO:210). Polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is thought to reside on chromosome 4. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 4.

This gene is expressed primarily in colon tissue, and to a lesser extent in thymus stromal cells and chondrosarcoma tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a

12

biological sample and for diagnosis of diseases and conditions which include, but are not limited to, chondroma and mannosidosis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the chondro and immune system. The expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, metabolic, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

10

15

20

25

30

The tissue distribution and homology to bovine beta-mannosidase indicates that the protein products of this gene are useful for the diagnosis and/or treatment of chondroma and mannosidosis. Human beta-mannosidosis is an autosomal recessive, lysosomal storage disease caused by a deficiency of the enzyme beta-mannosidase. Furthermore, the homology of the translation product of this gene to beta-mannosidase indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, prevention, and/or treatment of various metabolic disorders such as lysosomal storage deficiencies, Tay-Sachs disease, phenylkenonuria, galactosemia, hyperlipidemias, porphyrias, and Hurler's syndrome. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:13 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2093 of SEQ ID NO:13, b is an

integer of 15 to 2107, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:13, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 4

5

10

15

20

25

30

In specific embodiments, polypeptides of the invention comprise the following amino acid sequences: PRLTPRMKWPTAALASRLLGWTVLRPPYPRVPSLPOVT LHPTDGLMAVLYTGGEGRTLGEQHFFHETFVTRWLLGPVPVRFGACSPLSFL APRRGQGAPAGXFCACPRPASRQLCPWPALPGTPYSNSAPLCTGMGHSNTPQ GPPSPQYALSPTEPTSLSGNSHLPAILVL (SEO NO:211), PRLTPRMKWPTAAL ASRLLGWTVLRPPYPRVPSLPQVTLHP (SEQ ID NO:212), TDGLMAVLYTGGE GRTLGEQHFFHETFVTRWLLGPVPVRFG (SEQ ID NO:213), ACSPLSFLAPRRGQGAPAGXFCACPRPAS RQLCPWPALPGTP (SEQ I D NO:214), and/or YSNSAPLCTGMGHSNTPQGPPSPQYALSPTEPTSLSGNS HLPAILVL (SEQ ID NO:215). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in human lung (adult and fetal), and to a lesser extent in liver and brain tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, pulmonary disorders and hemostasis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the lung and liver tissues, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., pulmonary, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

14

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:

5

10

15

20

25

30

111 as residues: Arg-28 to Gln-36.

The tissue distribution in lung and liver tissues indicates that the protein products of this gene are useful for the diagnosis and/or treatment of pulmonary disorders and hematopoietic disorders. The tissue distribution in adult and fetal lung tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection and treatment of disorders associated with developing lungs, particularly in premature infants where the lungs are the last tissues to develop. The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and intervention of lung tumors, since the gene may be involved in the regulation of cell division, particularly since it is expressed in fetal tissue. Alternatively,

Expression of this gene product in liver tissue indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:14 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1248 of SEQ ID NO:14, b is an integer of 15 to 1262, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:14, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 5

5

10

15

20

25

In specific embodiments, polypeptides of the invention comprise the following amino acid sequences: HLLEVTPCRLPVPEFPGRTPRGSRTPD (SEQ ID NO:216). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in rapidly dividing liver tissue, (e.g., hepatoma, hepatocellular carcinoma, and fetal liver tissue), and to a lesser extent in normal liver tissue, and other tumors such as colon cancer and uterine cancer.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancers, particularly hepatomas, colon cancer, and uterine cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the liver, colon and uterus, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., liver, colon, uterus, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

5

10

15

20

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 112 as residues: Trp-35 to Trp-45, Pro-52 to Asp-57, Thr-73 to Arg-82, Pro-105 to Leu-112, Pro-115 to Arg-127, Pro-140 to Gln-151.

The tissue distribution in liver tissues and cancers thereof, as well as other cancerous tissues, indicates that the protein products of this gene are useful for the diagnosis and/or treatment of cancers, particularly, hepatoma, colon cancer and uterine cancer, as well as cancers of other tissues where expression has been observed. Furthermore, expression within cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Thus, this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:15 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 745 of SEQ ID NO:15, b is an integer of 15 to 759, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:15, and where b is greater than or equal to a + 14.

25

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 6

This gene is expressed primarily in hepatocellular tumors.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hepatomas. Similarly, polypeptides and antibodies directed to these

17

polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the liver, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., liver, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

5

10

15

20

25

30

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 113 as residues: Pro-32 to Gly-40.

The tissue distribution in hepatocellular tumors indicates that the protein products of this gene are useful for the diagnosis and/or treatment of hepatomas, as well as cancers of other tissues where expression has been observed. Furthermore, expression within cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Thus, this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:16 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1796 of SEQ ID NO:16, b is an integer of 15 to 1810, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:16, and where b is greater than or equal to a + 14.

18

FEATURES OF PROTEIN ENCODED BY GENE NO: 7

5

10

15

20

25

30

This gene is expressed primarily in human rhabdomyosarcoma tissue, as well as in placental tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, malignant neoplasms and reproductive disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skeletal system and reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., reproductive, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 114 as residues: Arg-23 to Trp-28, Phe-93 to Lys-98, Arg-199 to Trp-206, Gly-208 to Met-213.

The tissue distribution in placental tissue and human rhabdomyosarcoma tissue indicates that the protein products of this gene are useful for the diagnosis and/or treatment of skeletal and reproductive disorders. Furthermore, the tissue distribution in placental tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of disorders of the placenta. Specific expression within the placenta indicates that this gene product may play a role in the proper establishment and maintenance of placental function. Alternately, this gene product may be produced by the placenta and then transported to the embryo, where it may play a crucial role in the development and/or survival of the developing embryo or fetus.

Expression of this gene product in a vascular-rich tissue such as the placenta also indicates that this gene product may be produced more generally in endothelial cells or within the circulation. In such instances, it may play more generalized roles in vascular function, such as in angiogenesis. It may also be produced in the vasculature and have effects on other cells within the circulation, such as hematopoietic cells. It may serve to promote the proliferation, survival, activation, and/or differentiation of hematopoietic cells, as well as other cells throughout the body. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:17 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1038 of SEQ ID NO:17, b is an integer of 15 to 1052, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:17, and where b is greater than or equal to a + 14.

20

25

30

5

10

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 8

This gene is expressed primarily in fetal liver/spleen and fetal skin tissues, and to a lesser extent in breast cancer tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, developmental disorders and neoplasias. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the fetal tissue and adult immune system, expression of this gene at significantly higher or lower levels may be

20

routinely detected in certain tissues or cell types (e.g., developing, immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

5

10

15

20

25

30

The tissue distribution in fetal liver/spleen and skin tissues indicates that the protein products of this gene are useful for the diagnosis and/or treatment of developmental disorders and malignant neoplasias. Likewise, expression within fetal tissue and other cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, fetal development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus, this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy.

Alternatively, the tissue distribution in fetal skin tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment, diagnosis, and/or prevention of various skin disorders including congenital disorders (i.e. nevi, moles, freckles, Mongolian spots, hemangiomas, port-wine syndrome), integumentary tumors (i.e. keratoses, Bowen's disease, basal cell carcinoma, squamous cell carcinoma, malignant melanoma, Paget's disease, mycosis fungoides, and Kaposi's sarcoma), injuries and inflammation of the skin (i.e.wounds, rashes, prickly heat disorder, psoriasis, dermatitis), atherosclerosis, uticaria, eczema, photosensitivity, autoimmune disorders (i.e. lupus erythematosus, vitiligo, dermatomyositis, morphea, scleroderma, pemphigoid, and pemphigus), keloids, striae, erythema, petechiae, purpura, and xanthelasma. Moreover, such disorders may predispose increased susceptibility to viral and bacterial infections of the skin (i.e. cold sores, warts, chickenpox, molluscum contagiosum, herpes zoster, boils, cellulitis, erysipelas, impetigo, tinea, althletes foot, and ringworm). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

5

10

15

20

25

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:18 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1116 of SEQ ID NO:18, b is an integer of 15 to 1130, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:18, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 9

The translation product of this gene shares sequence homology with the bacterial gufA gene, as well as a C. elegans protein of unknown function.

In specific embodiments, polypeptides of the invention comprise the following amino acid sequences: MIPGSDSQTALNFGSTLMKKKSDPEGPALLFPESELSIRI GRAGLLSDKSENGEAYQRKKAAATGLPEGPAVPVPSRGNLAQPGGSSWRRI ALLILAITIHNVPEGLAVGVGFGAIEKTASATFESARNLAIGIGIONFPEGLAVS LPLRGAGFSTWRAFWYGQLSGMVEPLAGVFGAFAVVLAEPILPYALAFAAG AMVYVVMDDIIPEAQISGNGKLASWASILGFVVMMSLDVGLG (SEQ ID NO:217), MIPGSDSQTALNFGSTLMKKKSDPEGPALLFPESELSIRIGRA (SEQ ID NO:218), GLLSDKSENGEAYQRKKAAATGLPEGPAVPVPSRGNLAQPG (SEOI D NO: 219), GSSWRRIALLILAITIHNVPEGLAVGVGFGAIEKTASATFESAR (SEO ID NO:220), NLAIGIGIQNFPEGLAVSLPLRGAGFSTWRAFWYGQLS GMVEP (SEQ ID NO:221), LAGVFGAFAVVLAEPILPYALAFAAGAMVYVVM DDIIPEAQIS (SEQ ID NO:222), and/or GNGKLASWASILGFVVMMSLDVGLG (SEQ ID NO:223). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in cells of the immune system, particularly macrophage.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the immune system, such as AIDS, as well as inflammatory disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in immune cells such as macrophage indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Expression of this gene product in macrophage also strongly indicates a role

for this protein in immune function and immune surveillance. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

23

PCT/US99/05804

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:19 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 869 of SEQ ID NO:19, b is an integer of 15 to 883, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:19, and where b is greater than or equal to a + 14.

15 FEATURES OF PROTEIN ENCODED BY GENE NO: 10

5

10

20

25

30

This gene is expressed primarily in the spleen metastic melanoma tissue as well as in embryonic tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders affecting the spleen or immune system, developmental disorders, and cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., spleen, developing, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

5

10

15

20

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 117 as residues: Asn-37 to Lys-44, Ser-73 to Glu-78, Ala-103 to Ser-111.

The tissue distribution in spleen metastic melanoma and embryonic tissues indicates that the protein products of this gene are useful for the diagnosis and/or treatment of disorders affecting the spleen, including cancers of the spleen, as well as cancers of other tissues where expression has been observed. Furthermore, expression within embryonic tissue and other cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus, this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:20 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 975 of SEQ ID NO:20, b is an integer of 15 to 989, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:20, and where b is greater than or equal to a + 14.

25

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 11

It has been discovered that this gene is expressed primarily in cells of the immune system, including monocytes and neutrophils.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: disorders affecting the immune

10

15

20

25

30

system, such as AIDS. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 118 as residues: Ser-12 to Asp-20, Gly-22 to Gly-32, Ala-49 to Thr-57.

The tissue distribution in monocytes and neutrophils indicates that the protein products of this clone are useful for the diagnosis and/or treatment of immune system disorders, including AIDS. Furthermore, expression of this gene product in monocytes and neutrophils suggests a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Expression of this gene product in monocytes and neutrophils also strongly suggests a role for this protein in immune function and immune surveillance. Protein,

as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:21 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 481 of SEQ ID NO:21, b is an integer of 15 to 495, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:21, and where b is greater than or equal to a + 14.

15 FEATURES OF PROTEIN ENCODED BY GENE NO: 12

5

10

20

25

30

It has been discovered that this gene is expressed primarily in cells of the immune system, including monocytes.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: disorders affecting the immune system. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 119 as residues: Glu-35 to Trp-42.

The tissue distribution suggests that the protein product of this clone is useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in monocytes suggests a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Expression of this gene product in monocytes also strongly suggests a role for this protein in immune function and immune surveillance. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:22 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2303 of SEQ ID NO:22, b is an integer of 15 to 2317, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:22, and where b is greater than or equal to a + 14.

5

10

15

20

25

30

WO 99/47540 PCT/US99/05804

28

FEATURES OF PROTEIN ENCODED BY GENE NO: 13

It has been discovered that this gene is expressed primarily in cells of the immune system, including monocytes.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: disorders of the immune system. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

The tissue distribution in monocytes indicates that the protein products of this clone are useful for the diagnosis and/or treatment of disorders of the immune system. Expression of this gene product in monocytes suggests a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of

various blood lineages, and in the differentiation and/or proliferation of various cell types. Expression of this gene product in monocytes also strongly suggests a role for this protein in immune function and immune surveillance. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:23 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1712 of SEQ ID NO:23, b is an integer of 15 to 1726, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:23, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 14

5

10

15

20

30

The translation product of this gene shares sequence homology with a gene from C. elegans of unknown function.

In specific embodiments, polypeptides of the invention comprise the following amino acid sequences: TRPITYVLLAG (SEQ ID NO:224). Polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is thought to reside on chromosome 11. Accordingly,

25 polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 11.

It has been discovered that this gene is expressed primarily in fetal lung, liver, spleen and heart tissues, as well as adult liver, bladder, endometrial stromal cells, synovium, colon cancer, smooth muscle, keratinocytes, and the bone marrow derived cell line RS4;11.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: disorders of the musculo-skeletal system, and cancers of the immune system. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the musculo-skeletal and immune systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., immune, musculo-skeletal, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

The tissue distribution in tissues of the immune system indicates that the protein products of this clone are useful for treating proliferative disorders of immune system precursor cells. Alternatively, the tissue distribution in smooth muscle and heart tissue indicates that the protein product of this gene is useful for the diagnosis and treatment of conditions and pathologies of the cardiovascular system, such as heart disease, restenosis, atherosclerosis, stoke, angina, thrombosis, and wound healing. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:24 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 515 of SEQ ID NO:24, b is an integer of 15 to 529, where both a and b correspond to the positions of

31

PCT/US99/05804

nucleotide residues shown in SEQ ID NO:24, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 15

5

10

15

20

25

30

In specific embodiments, polypeptides of the invention comprise the following amino acid sequences: GTSLTAPLLEFLLALYFLFADAMQLNDKWQGLCWP (SEQ ID NO:225). Polynucleotides encoding these polypeptides are also encompassed by the invention.

It has been discovered that this gene is expressed primarily in T-cells, fetal spleen and infant brain tissues, and to a lesser extent in many other tissues including melanocytes, lung cancer, macrophages, dendritic cells, stromal cells, adrenal gland and others.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: inflammation and autoimmunity, developing tissues. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and developing system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., immune, developing, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 122 as residues: Ser-46 to Gly-51.

The tissue distribution in T-cells and other immune cells indicates that the protein products of this clone are useful for treating diseases involving the activation of T-cells, including inflammation and autoimmune diseases. Alternatively, the tissue distribution in a wide range of fetal tissues suggests that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and

treatment of cancer and other proliferative disorders. Similarly, fetal development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus, this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

5

10

15

30

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:25 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1741 of SEQ ID NO:25, b is an integer of 15 to 1755, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:25, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 16

In specific embodiments, polypeptides of the invention comprise the following amino acid sequences: LANFZCSDCAQTVLFVLZFZILVFTYEIPF (SEQ ID NO:226). Polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is thought to reside on chromosome 13. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 13. Recently another group published this gene, referring to it as CLN5 (See Genbank Accession No.: 3342386).

It has been discovered that this gene is expressed primarily in placental tissue, 12 week embryos, and tumors including testes, tongue and pharynx, and to a lesser extent in adipose tissue, tonsils, melanocytes, fetal spleen, macrophages, T-cells, amniotic cells, and brain tissue.

5

10

15

20

25

30

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: tumors, particularly of the tongue and throat, and neurodegenerative disorders. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the neural and digestive systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., tongue, throat, brain, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 123 as residues: Pro-44 to Ala-60, Val-187 to Thr-193, Lys-203 to Ala-210, Thr-212 to Cys-219.

The tissue distribution in tongue and pharynx carcinoma tissue indicates that the protein products of this clone are useful for diagnosing and/or treating oral cancers, including tumors of the throat and tongue. Furthermore, the tissue distribution in brain tissue suggests that the protein product of this clone is useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as neuronal ceroid lipofuscinoses (NCLs), Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, or sexually-linked disorders. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are

PCT/US99/05804

34

related to SEQ ID NO:26 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1737 of SEQ ID NO:26, b is an integer of 15 to 1751, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:26, and where b is greater than or equal to a + 14.

10

20

25

30

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 17

In specific embodiments, polypeptides of the invention comprise the following amino acid sequences:

QAWHEVGGGVRRCWFVLGERRAGSLLSASYGTFAMPG

MVLFGRRWAIASDDLVFPGFFELVVRVLWWIGILTLYL (SEQ ID NO:227), and/or PGMVLFGRRWAIASDDLVFPGFFELVVRVLWWIGILTLYLMHRGKLD CAGGALLSSYLIVLMILLAVVICTVSAIMCVSMRGTICNPGPRKSMSKLLYIRL ALFFPEMVWASLGAAWVADGVQCD (SEQ ID NO:228). Polynucleotides encoding these polypeptides are also encompassed by the invention.

It has been discovered that this gene is expressed in activated neutrophils, infant brain tissue and primary dendritic cells.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: disorders of the immune system, and neurodegenerative disorders. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and neural systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., immune, brain, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual

35

having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

5

10

15

20

25

30

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 124 as residues: Pro-47 to Met-53, Ser-130 to Ser-138.

The tissue distribution in neutrophils and primary dendritic cells indicates that the protein products of this clone are useful for diagnosing and/or treating immune system disorders. Expression of this gene product in neutrophils and primary dendritic cells suggests a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Expression of this gene product in neutrophils and primary dendritic cells also strongly suggests a role for this protein in immune function and immune surveillance.

Alternatively, the tissue distribution in brain tissue suggests that the protein product of this clone is useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, or sexually-linked disorders. Protein, as well as, antibodies

10

20

25

30

directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:27 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1198 of SEQ ID NO:27, b is an integer of 15 to 1212, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:27, and where b is greater than or equal to a + 14.

15 FEATURES OF PROTEIN ENCODED BY GENE NO: 18

It has been discovered that this gene is expressed primarily in neutrophils, and to a lesser extent in other tissues.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: immune and inflammatory disorders. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 125 as residues: Gln-17 to Ser-24.

WO 99/47540 PCT/US99/05804

The tissue distribution in neutrophils indicates that the protein products of this clone are useful for the diagnosis and/or treatment of immune and inflammatory disorders. Expression of this gene product in neutrophils suggests a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Expression of this gene product in neutrophils also strongly suggests a role for this protein in immune function and immune surveillance. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:28 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1098 of SEQ ID NO:28, b is an integer of 15 to 1112, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:28, and where b is greater than or equal to a + 14.

38

FEATURES OF PROTEIN ENCODED BY GENE NO: 19

5

10

15

20

25

30

In specific embodiments, polypeptides of the invention comprise the following amino acid sequences: HERNCFPMWLNHSAFPPV (SEQ ID NO:229).

Polynucleotides encoding these polypeptides are also encompassed by the invention.

It has been discovered that this gene is expressed primarily in neutrophils, and to a lesser extent in other tissues.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: immune and inflammatory disorders. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

The tissue distribution in neutrophils indicates that the protein products of this clone are useful for the diagnosis and/or treatment of immune and inflammatory disorders. Expression of this gene product in neutrophils suggests a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune

deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Expression of this gene product in neutrophils also strongly suggests a role for this protein in immune function and immune surveillance. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:29 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 734 of SEQ ID NO:29, b is an integer of 15 to 748, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:29, and where b is greater than or equal to a + 14.

20

10

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 20

In specific embodiments, polypeptides of the invention comprise the following amino acid sequences: GWTRENDHRALSKAGIGSAEIQPSNLRVGSAKDLGKPW AGKLLLLSSCLLFFSLGVLYRGQMLAPPLQEDWKGGVKDSDLIDDSSASPIPP SYLEYKAALYPFSEHKSVRNATDSLTFFLVTDHFLDNQDSQ (SEQ ID NO:230), GWTRENDHRALSKAGIGSAEIQPSNLRVGSAKDLGKPWAGKLLLL (SEQ ID NO:231), SSCLLFFSLGVLYRGQMLAPPLQEDWKGGVKDSDLIDDSSASPIPP (SEQ ID NO:232), and/or SYLEYKAALYPFSEHKSVRNATDSLTFFLVTDHFL DNQDSQ (SEQ ID NO:233). Polynucleotides encoding these polypeptides are also encompassed by the invention.

10

15

20

25

30

It has been discovered that this gene is expressed primarily in ovarian cancer tissue, and to a lesser extent in other tissues.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: ovarian cancer. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the ovaries, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., reproductive, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 127 as residues: Thr-20 to Gly-27, Gly-32 to Phe-41.

The tissue distribution in ovarian cancer tissue indicates that the protein products of this clone are useful for the diagnosis and/or treatment of ovarian cancer, as well as cancers of other tissues where expression has been observed. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:30 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 764 of SEQ ID NO:30, b is an integer of 15 to 778, where both a and b correspond to the positions of

41

nucleotide residues shown in SEQ ID NO:30, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 21

5

10

15

20

25

30

When tested against U937 Myeloid cell lines, supernatants removed from cells containing this gene activated the GAS assay. Thus, it is likely that this gene activates myeloid cells, and to a lesser extent other cells, through the Jak-STAT signal transduction pathway. The gamma activating sequence (GAS) is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

In specific embodiments, polypeptides of the invention comprise the following amino acid sequences: LKFHQESLSGD (SEQ ID NO:234). Polynucleotides encoding these polypeptides are also encompassed by the invention.

It has been discovered that this gene is expressed primarily in fast-growing tissues such as immune/hematopoietic tissues, early developmental stage human tissues, and tumor tissues, and to a lesser extent in some other tissues.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: growth disorders, immune and inflammatory diseases, and tumorigenesis. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune/hematopoietic system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

10

15

20

25

30

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 128 as residues: Glu-60 to Arg-65.

The tissue distribution in immune tissues, in conjunction with the biological activity data, indicates that the protein products of this clone are useful for the diagnosis and/or treatment of growth disorders, immune and inflammatory diseases, and tumorigenesis. Furthermore, expression within embryonic tissue and other cellular sources marked by proliferating cells suggests that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus, this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:31 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1310 of SEQ ID NO:31, b is an integer of 15 to 1324, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:31, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 22

In specific embodiments, polypeptides of the invention comprise the following amino acid sequences: EAKSRPVTQAGVQWHDLGSLQPLPP (SEQ ID NO:235). Polynucleotides encoding these polypeptides are also encompassed by the invention.

5

10

15

20

25

30

43

It has been discovered that this gene is expressed primarily in ovarian cancer tissue, and to a lesser extent in fetal liver/spleen and retinal tissues.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: ovarian cancer, immune disorders, and retinal disorders. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the ovaries, immune and ocular systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., reproductive, ovaries, retina, immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

The tissue distribution in ovarian cancer tissue indicates that the protein products of this clone are useful for the diagnosis and/or treatment of ovarian cancer, as well as cancers of other tissues where expression has been observed. The tissue distribution also suggests that the protein product of this clone is useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in fetal liver/spleen suggests a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have

PCT/US99/05804

commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Alternatively, the tissue distribution in retinal tissue suggests that the protein product of this clone is useful for the treatment and/or detection of eye disorders including blindness, color blindness, impaired vision, short and long sightedness, retinitis pigmentosa, retinitis proliferans, and retinoblastoma, retinochoroiditis, retinopathy and retinoschisis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:32 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 725 of SEQ ID NO:32, b is an integer of 15 to 739, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:32, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 23

10

15

20

25

30

The translation product of this gene shares sequence homology with a C. elegans protein of unknown function (See Genbank Accession No.: gnllPIDle1348017). When tested against fibroblast cell lines, supernatants removed from cells containing this gene activated the EGR1 assay. Thus, it is likely that this gene activates fibroblast cells through a signal transduction pathway. Early growth response 1 (EGR1) is a promoter associated with certain genes that induces various tissues and cell types upon activation, leading the cells to undergo differentiation and proliferation. The gene encoding the disclosed cDNA is thought to reside on

chromosome 17. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 17.

In specific embodiments, polypeptides of the invention comprise the following amino acid sequences: EAKSRPVTQAGVQWHDLGSLQPLPP (SEQ ID NO:236), and/or ALVLVCRQRYCRPRDLLQRYDSKPIVDLIGAMETQSEPSELELDDVVIT NPHIEAILENEDWIEDASGLMSHCIAILKICHTLTEKLVAMTMGSGAKMKTSA SVSDIIVVAKRISPRVDDVVKSMYPPLDPKLLDAR (SEQ ID NO:237).

5

10

15

20

25

Polynucleotides encoding these polypeptides are also encompassed by the invention.

It has been discovered that this gene is expressed primarily in fast growing tissues such as early development stage human tissues, immune/hematopoietic tissues, melanocytes, and tumor tissues, and to a lesser extent in some other tissues.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: growth disorders, immune and inflammatory disoders, skin and connective tissue disorders, and tumorigenesis. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the fast growing tissues such as early development stage human tissues,

immune/hematopoietic tissues, skin and connective tissue, and tumor tissues, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., musculo-skeletal, skin, immune, developing, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 130 as residues: Pro-34 to Ser-43, Glu-54 to Ser-60.

The tissue distribution suggests that the protein product of this clone is useful for the diagnosis and/or treatment of growth disorders, immune and inflammatory disorders, and tumorigenesis. Alternatively, the tissue distribution in melanocytes, in

46

conjunction with the observed biological activity data, suggests that the protein product of this clone is useful for the treatment, diagnosis, and/or prevention of various skin disorders including congenital disorders (i.e. nevi, moles, freckles, Mongolian spots, hemangiomas, port-wine syndrome), integumentary tumors (i.e. keratoses, Bowen's disease, basal cell carcinoma, squamous cell carcinoma, malignant melanoma, Paget's disease, mycosis fungoides, and Kaposi's sarcoma), injuries and inflammation of the skin (i.e.wounds, rashes, prickly heat disorder, psoriasis, dermatitis), atherosclerosis, uticaria, eczema, photosensitivity, autoimmune disorders (i.e. lupus erythematosus, vitiligo, dermatomyositis, morphea, scleroderma, pemphigoid, and pemphigus), keloids, striae, erythema, petechiae, purpura, and xanthelasma.

Moreover, such disorders may predispose increased susceptibility to viral and bacterial infections of the skin (i.e. cold sores, warts, chickenpox, molluscum contagiosum, herpes zoster, boils, cellulitis, erysipelas, impetigo, tinea, althletes foot, and ringworm). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and immunotherapy targets for the above listed tumors and tissues.

15

20

25

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:33 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1448 of SEQ ID NO:33, b is an integer of 15 to 1462, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:33, and where b is greater than or equal to a + 14.

10

FEATURES OF PROTEIN ENCODED BY GENE NO: 24

When tested against U937 Myeloid cell lines, supernatants removed from cells containing this gene activated the GAS assay. Thus, it is likely that this gene activates myeloid cells, and to a lesser extent other cells, through the Jak-STAT signal transduction pathway. The gamma activating sequence (GAS) is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells

In specific embodiments, polypeptides of the invention comprise the following amino acid sequences: DVESRGPSARCLPVVPGSLLPGLEPATKLMPGGLAPGHG APVRELLLPLLSQPTLGSLWDSLRHCSLLCNPLSCVPALEAPPSLVSLGCSGGC PRLSLAGSASPFPFLTALLSLLNTLAQIHKGLCGOLAAILAAPGLONYFLOCVA 15 PGAAPHLTPFSAWALRHEYHLQYLALALAQKAAALQPLPATHAALYHGMAL ALLSRLLPGSEYLTHELLLSCVFRLEFLPERTSGGPEAADFSDOLSLGSSRVPR CGQGTLLAQACQDLPSIRNCYLTHCSPARASLLASQALHRGELQRVPTLLLP MPTEPLLPTDWPFLH (SEQ I D DVESRGPSARCLPVVPGSLLPGLEPATKLM PGGLAPGHGAPVRE (SEO ID 20 NO:239), LLLPLLSQPTLGSLWDSLRHCSLLCNP LSCVPALEAPPSLVSLGC ID (SEO NO:240), SGGCPRLSLAGSASPFPFLTALL SLLNTLAQIHKGLCGQLAAILA (SEQ ID NO:241), APGLONYFLQCVAPGAAP HLTPFSAWALRHEYHLQYLALALAQK (SEO ID NO:242), AAALOPLPATHAA 25 LYHGMALALLSRLLPGSEYLTHELLLSCVFR (SEQ ID NO:243), LEFLPERTSG GPEAADFSDQLSLGSSRVPRCGQGTLLAQACQDL (SEQ ID NO:244), and/or PSIRNCYLTHCSPARASLLASQALHRGELQRVPTLLLPMPTEPLLPTDWPFLH (SEO ID NO:245). Polynucleotides encoding these polypeptides are also encompassed by the invention.

10

15

20

25

30

It has been discovered that this gene is expressed primarily in hematopoietic tissues and fetal heart tissue, and to a lesser extent in brain and gall bladder tissues, and some other tissues.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: immune and inflammatory disorders, cardiovascular disorders, and growth disorders. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hematopoietic and vascular systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., vascular, immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 131 as residues: Tyr-88 to Trp-102, Asp-105 to Ser-110.

The tissue distribution in hematopoietic tissues, in conjunction with the observed biological activity data, indicates that the protein products of this clone are useful for the diagnosis and/or treatment of immune and inflammatory disorders and growth disorders. Alternatively, the tissue distribution in fetal heart tissue indicates that the protein product of this gene is useful for the diagnosis and treatment of conditions and pathologies of the cardiovascular system, such as heart disease, restenosis, atherosclerosis, stoke, angina, thrombosis, and wound healing. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:34 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically

49

excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2801 of SEQ ID NO:34, b is an integer of 15 to 2815, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:34, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 25

5

10

15

20

25

30

In specific embodiments, polypeptides of the invention comprise the following amino acid sequences: VGSVLGAFLTFPGLRLAQTHRDALT (SEQ ID NO:246). Polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is thought to reside on chromosome 19. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 19.

It has been discovered that this gene is expressed primarily in human pituitary tissue.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: hyperpituitarism and hypopituitarism. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., endocrine, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder. This gene is found on the short arm of chromosome 19 and, therefore, is useful as a chromosome marker.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 132 as residues: Met-1 to Pro-6, Gln-89 to Ala-94, Pro-161 to Cys-173.

The tissue distribution in pituitary tissue indicates that the protein products of this clone are useful for the diagnosis and/or treatment of pituitary disorders. More generally, the tissue distribution in pituitary tissue suggests that the protein product of this clone is useful for the detection, treatment, and/or prevention of various endocrine disorders and cancers, particularly Addison's disease, Cushing's Syndrome, and disorders and/or cancers of the pancrease (e.g. diabetes mellitus), adrenal cortex, ovaries, pituitary (e.g., hyper-, hypopituitarism), thyroid (e.g. hyper-, hypothyroidism), parathyroid (e.g. hyper-, hypoparathyroidism), hypothallamus, and testes. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

5

10

15

20

30

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:35 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1064 of SEQ ID NO:35, b is an integer of 15 to 1078, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:35, and where b is greater than or equal to a + 14.

25 FEATURES OF PROTEIN ENCODED BY GENE NO: 26

It has been discovered that this gene is expressed highly and specifically in placental and bone marrow cDNA libraries, and to a lesser extent in T-cells.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: immune, developmental and reproductive disorders. Similarly, polypeptides and antibodies directed to those

WO 99/47540

10

15

20

25

30

polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and developing systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., immune, developmental, reproductive, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

The tissue distribution in bone marrow and placental tissue indicates that the protein products of this clone are useful for the diagnosis and/or treatment of immune and reproductive disorders. The tissue distribution in bone marrow suggests that the protein product of this clone is useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia. The uses include bone marrow cell ex vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types.

Alternatively, the tissue distribution in placental tissue suggests that the protein product of this clone is useful for the diagnosis and/or treatment of disorders of the placenta. Specific expression within the placenta suggests that this gene product may play a role in the proper establishment and maintenance of placental function. Alternately, this gene product may be produced by the placenta and then transported to the embryo, where it may play a crucial role in the development and/or survival of the developing embryo or fetus.

Expression of this gene product in a vascular-rich tissue such as the placenta also suggests that this gene product may be produced more generally in endothelial cells or within the circulation. In such instances, it may play more generalized roles in

10

15

25

30

vascular function, such as in angiogenesis. It may also be produced in the vasculature and have effects on other cells within the circulation, such as hematopoietic cells. It may serve to promote the proliferation, survival, activation, and/or differentiation of hematopoietic cells, as well as other cells throughout the body. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:36 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1203 of SEQ ID NO:36, b is an integer of 15 to 1217, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:36, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 27

In specific embodiments, polypeptides of the invention comprise the following amino acid sequences:

LECTDTIMVHCSLKLLSPSDXSHSASQVAKTRGVHHXTQ LIFKVFFVXMGSHSTKYXSIRPGLLP (SEQ ID NO:247). Polynucleotides encoding these polypeptides are also encompassed by the invention.

It has been discovered that this gene is expressed primarily in human prostate and smooth muscle tissues.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: disorders in the prostate gland, vascular and connective tissues. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential

identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the male reproductive and urinary system and vascular system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., reproductive, vascular, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

The tissue distribution in prostate and smooth muscle tissues indicates that the protein products of this clone are useful for the diagnosis and/or treatment of prostate gland, vascular and connective tissue disorders. The tissue distribution in smooth muscle tissue indicates that the protein product of this gene is useful for the diagnosis and treatment of conditions and pathologies of the cardiovascular system, such as heart disease, restenosis, atherosclerosis, stoke, angina, thrombosis, and wound healing. The expression in the prostate tissue may indicate the gene or its products can be used in the disorders of the prostate, including inflammatory disorders, such as chronic prostatitis, granulomatous prostatitis and malacoplakia, prostatic hyperplasia and prostate neoplastic disorders, including adenocarcinoma, transitional cell carcinomas, ductal carcinomas, squamous cell carcinomas, or as hormones or factors with systemic or reproductive functions. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:37 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1268 of SEQ ID NO:37, b is an integer of 15 to 1282, where both a and b correspond to the positions of

54

nucleotide residues shown in SEQ ID NO:37, and where b is greater than or equal to a + 14.

10

15

20

25

30

WO 99/47540 PCT/US99/05804

55

FEATURES OF PROTEIN ENCODED BY GENE NO: 28

In specific embodiments, polypeptides of the invention comprise the following amino acid sequences: ESSFVPPAAHSSLC (SEQ ID NO:248). Polynucleotides encoding these polypeptides are also encompassed by the invention.

It has been discovered that this gene is expressed primarily in human pituitary tissue.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: hyperpituitarism and hypopituitarism. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., endocrine, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

The tissue distribution in pituitary tissue indicates that the protein products of this clone are useful for the diagnosis and/or treatment of pituitary gland disorders such as hypopituitarism and hyperpituitarism. More generally, the tissue distribution in pituitary tissue suggests that the protein product of this clone is useful for the detection, treatment, and/or prevention of various endocrine disorders and cancers, particularly Addison's disease, Cushing's Syndrome, and disorders and/or cancers of the pancrease (e.g. diabetes mellitus), adrenal cortex, ovaries, pituitary (e.g., hyper-, hypopituitarism), thyroid (e.g. hyper-, hypothyroidism), parathyroid (e.g. hyper-, hypoparathyroidism), hypothallamus, and testes. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are

WO 99/47540

related to SEQ ID NO:38 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 545 of SEQ ID NO:38, b is an integer of 15 to 559, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:38, and where b is greater than or equal to a + 14.

10

25

30

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 29

In specific embodiments, polypeptides of the invention comprise the following amino acid sequences: LLPGQQEATQCVEAGAGEGALTPMCPCRQEQFVDLYKEF 15 EPSLVNSTVYIMAMAIQMAPFAINYKVRPGPCXNIHCLPTQPHPMKPSVPHPH RARPSWRACPRTSPWCGVWQFHSWPSLACSSAPRPTSTASLASWTSLWSSS WSLPRSCSWTSAWRSWPTASCSSSWGPRS (SEO ID NO:249), LLPGQQEATQCV EAGAGEGALTPMCPCRQEQFVDLYKEFEPSLVN (SEQ ID NO:250), STVYIMAMAIQMAPFAINYKVRPGPCXNIHCLPTQPHPMKPSVP 20 (SEQ I D N O : 2 5 1) ,HPHRARPSWRACPRTSPWCGVWQFHSWPSLACSSAPRPTSTA (SEO ID NO:252), and/or SLASWTSLWSSSWSLPRSCSWTSAWRSWPTASCSSSWG PRS (SEQ ID NO:253). Polynucleotides encoding these polypeptides are also encompassed by the invention.

It has been discovered that this gene is expressed primarily in human pituitary and breast tissues, and to a lesser extent in endometrial and ovarian cancer tissues.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: hyperpituitarism and hypopituitarism, and cancers of the female reproductive system. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide

WO 99/47540

5

10

15

20

25

30

57

PCT/US99/05804

immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine and reproductive systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., endocrine, reproductive, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 136 as residues: Ser-3 to Lys-8.

The tissue distribution in pituitary tissue indicates that the protein products of this clone are useful for the diagnosis and/or treatment of disorders in the pituitary gland. More generally, the tissue distribution in pituitary tissue suggests that the protein product of this clone is useful for the detection, treatment, and/or prevention of various endocrine disorders and cancers, particularly Addison's disease, Cushing's Syndrome, and disorders and/or cancers of the pancrease (e.g. diabetes mellitus), adrenal cortex, ovaries, pituitary (e.g., hyper-, hypopituitarism), thyroid (e.g. hyper-, hypothyroidism), parathyroid (e.g. hyper-, hypoparathyroidism), hypothallamus, and testes. Alternatively, the tissue distribution in breast tissue and cancerous tissues of the endometrium and ovaries suggests that the translation product of this gene is useful for the detection and/or treatment of disorders and cancers of the female reproductive system, as well as cancers of other tissues where expression has been observed. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:39 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the

general formula of a-b, where a is any integer between 1 to 789 of SEQ ID NO:39, b is an integer of 15 to 803, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:39, and where b is greater than or equal to a + 14.

5

10

15

20

25

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 30

In specific embodiments, polypeptides of the invention comprise the following amino acid sequences: TRNILSFIKCVIHNFWIPKESNEITIIINPYRETVCFSVEP VKKIFNY (SEQ ID NO:254). Polynucleotides encoding these polypeptides are also encompassed by the invention.

It has been discovered that this gene is expressed primarily in human synovial sarcoma tissue.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skeletal system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., skeletal, connective, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 137 as residues: Thr-29 to Pro-34.

The tissue distribution in synovial sarcoma tissue indicates that the protein products of this clone are useful for the diagnosis and/or treatment of diseases of the synovium. In addition, the

Expression of this gene product in synovium suggests a role in the detection and treatment of disorders and conditions affecting the skeletal system, in particular osteoporosis as well as disorders afflicting connective tissues (e.g. arthritis, trauma,

10

15

25

30

tendonitis, chrondomalacia and inflammation), such as in the diagnosis or treatment of various autoimmune disorders such as rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism, spinal deformation, and specific joint abnormalities as well as chondrodysplasias (ie. spondyloepiphyseal dysplasia congenita, familial arthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:40 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1496 of SEQ ID NO:40, b is an integer of 15 to 1510, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:40, and where b is greater than or equal to a + 14.

20 FEATURES OF PROTEIN ENCODED BY GENE NO: 31

In specific embodiments, polypeptides of the invention comprise the following amino acid sequences: LVVLFASSNSRYLKYFFLVPLILGSAW (SEQ ID NO:255). Polynucleotides encoding these polypeptides are also encompassed by the invention.

It has been discovered that this gene is expressed primarily in human rhabdomyosarcoma and fetal liver/spleen tissues.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: malignant neoplasms and hematopoiesis. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells.

60

particularly of the skeletal and immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., musculo-skeletal, immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 138 as residues: Gly-29 to Thr-35.

The tissue distribution in rhabdomyosarcoma and fetal liver/spleen tissues indicates that the protein products of this clone are useful for diagnosis and treatment of skeletal and immune disorders. The expression in rhabdomyosarcoma tissue suggests that the protein product of this clone is useful for the detection, treatment, and/or prevention of various muscle disorders, such as muscular dystrophy, cardiomyopathy, fibroids, myomas, and rhabdomyosarcomas. Alternatively,

10

15

20

25

30

Expression of this gene product in fetal liver/spleen tissue suggests a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

10

15

20

25

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:41 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1081 of SEQ ID NO:41, b is an integer of 15 to 1095, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:41, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 32

It has been discovered that this gene is expressed primarily in fibrosarcoma tissue.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: fibrosarcoma. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the connective tissue system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., musculo-skeletal, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 139 as residues: Ser-34 to Gln-40, Gly-42 to Glu-48, Tyr-56 to Leu-62.

The tissue distribution in only fibrosarcoma tissue suggests that the protein product of this clone is useful for the treatment, diagnosis and/or prognosis of

10

15

25

30

fibrosarcoma's or other diorders related with fibrous tissue including fibroma, fibromatosis, fibromyoma, fibromyositis, fibrosis and fibrositis. Likewise, the expression in fibrosarcoma tissue suggests that the protein product of this clone is useful for the detection, treatment, and/or prevention of various muscle disorders, such as muscular dystrophy, cardiomyopathy, myomas, and rhabdomyosarcomas. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:42 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1148 of SEQ ID NO:42, b is an integer of 15 to 1162, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:42, and where b is greater than or equal to a + 14.

20 FEATURES OF PROTEIN ENCODED BY GENE NO: 33

It has been discovered that this gene is expressed primarily in Hodgkins lymphoma and breast cancer tissues, and to a lesser extent in stromal cells and brain tissue.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: lymphoma, breast cancer, and neurological disorders. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune amd nervous systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types

10

15

20

25

PCT/US99/05804

(e.g., immune, neural, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 140 as residues: Pro-22 to Lys-29.

The tissue distribution in Hodgkins lymphoma, brain and breast cancer tissues suggests a role in the treatment, diagnosis and/or prognosis of breast cancer, immune and hematopoietic disorders including arthritis, asthma, immunodeficiency diseases, leukemia and Hodgkin's lymphoma and neurodegenerative disease states and behavioral disorders such as Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder and panic disorder. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:43 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 643 of SEQ ID NO:43, b is an integer of 15 to 657, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:43, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 34

In specific embodiments, polypeptides of the invention comprise the following
amino acid sequences: HEWKCKQKYSEGSGNTRIGN (SEQ ID NO:256).

Polynucleotides encoding these polypeptides are also encompassed by the invention.

64

It has been discovered that this gene is expressed primarily in chronic synovitis tissue, and to a lesser extent in fetal kidney and testes tissues.

5

10

15

20

25

30

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: synovitis, renal disorders and male infertility. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the connective tissue system, the renal system, and the male reproductive system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., skeletal, renal, reproductive, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 141 as residues: Met-33 to Pro-39, Ser-74 to Trp-79.

The tissue distribution of this gene in chronic synovitis, testes, and kidneys suggests a role in the treatment, diagnosis and prognosis of synovial membrane disorders including synovitis, renal disorders including kidney failure, renal colic, renal diabetes, hypertension, osteodystrophy, tubular acidosis and kidney stones: and and male infertility. Furthermore, the tissue distribution in testes tissue indicates that the protein product of this clone is useful for the treatment and/or diagnosis of conditions concerning proper testicular function (e.g. endocrine function, sperm maturation), as well as cancer. Therefore, this gene product is useful in the treatment of male infertility and/or impotence. This gene product is also useful in assays designed to identify binding agents, as such agents (antagonists) are useful as male contraceptive agents. Similarly, the protein is believed to be useful in the treatment and/or diagnosis of testicular cancer. The testes are also a site of active gene expression of transcripts that may be expressed, particularly at low levels, in other tissues of the body. Therefore, this gene product may be expressed in other specific

10

15

20

25

tissues or organs where it may play related functional roles in other processes, such as hematopoiesis, inflammation, bone formation, and kidney function, to name a few possible target indications. In addition, the

Expression of this gene product in synovium suggests a role in the detection and/or treatment of disorders and conditions affecting the skeletal system, in particular osteoporosis as well as disorders afflicting connective tissues (e.g. arthritis, trauma, tendonitis, chrondomalacia and inflammation), such as in the diagnosis or treatment of various autoimmune disorders such as rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism, spinal deformation, and specific joint abnormalities as well as chondrodysplasias (ie. spondyloepiphyseal dysplasia congenita, familial arthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:44 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1141 of SEQ ID NO:44, b is an integer of 15 to 1155, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:44, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 35

In specific embodiments, polypeptides of the invention comprise the following amino acid sequences: LLPLCFLGPRQVLEEFPSIV (SEQ ID NO:257).

30 Polynucleotides encoding these polypeptides are also encompassed by the invention.

66

It has been discovered that this gene is expressed primarily in brain tissue, and to a lesser extent in osteoclastoma and testes tissues.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: neurological disorders and male reproductive disorders. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous system and the male reproductive system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., neural, reproductive, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

5

10

15

20

25

30

The tissue distribution of this gene in brain tissue suggests a role in the diagnosis, prognosis and/or treatment of neurodegenerative disease states and behavioural disorders such as Alzheimer's Disease, Parkinson's Disease, Huntinton's Disease, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder and panic disorder. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, or sexually-linked disorders. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:45 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1098 of SEQ ID NO:45, b

PCT/US99/05804

is an integer of 15 to 1112, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:45, and where b is greater than or equal to a + 14.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 36

10

15

20

25

30

In specific embodiments, polypeptides of the invention comprise the following amino acid sequences: PTRPSKHQEAGS (SEQ ID NO:258). Polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is thought to reside on chromosome 3. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 3.

It has been discovered that this gene is expressed primarily in adult and fetal heart tissue, and to a lesser extent in fetal lung and fetal liver/spleen tissues.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: cardiovascular and immune disorders. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the vascular and immune systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., vascular, immune, pulmonary, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 143 as residues: Val-2 to Ser-14.

The tissue distribution in heart, fetal liver and fetal spleen tissues suggests a role in the treatment and/or diagnosis of cardiovascular disorders including myocardial infarction, congestive heart failure, coronary failure, as well as immune disorders including autoimmune diseases, such as lupus, transplant rejection, allergic

10

15

25

30

reactions, arthritis, asthma, immunodeficiency diseases, leukemia, and AIDS. Furthermore, the tissue distribution in adult and fetal heart tissue indicates that the protein product of this gene is useful for the diagnosis and treatment of conditions and pathologies of the cardiovascular system, such as heart disease, restenosis, atherosclerosis, stoke, angina, thrombosis, and wound healing. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:46 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 4009 of SEQ ID NO:46, b is an integer of 15 to 4023, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:46, and where b is greater than or equal to a + 14.

20 FEATURES OF PROTEIN ENCODED BY GENE NO: 37

It has been discovered that this gene is expressed primarily in testes tissues.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: male infertility and reproductive disorders. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the male reproductive system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., reproductive, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the

69

standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

The tissue distribution in testes tissues suggests a role in the treatment and/or diagnosis of male infertility, and testicular disorders including cancer. Furthermore, the tissue distribution in testes tissue indicates that the protein product of this clone is useful for the treatment and diagnosis of conditions concerning proper testicular function (e.g. endocrine function, sperm maturation), as well as cancer. Therefore, this gene product is useful in the treatment of male infertility and/or impotence. This gene product is also useful in assays designed to identify binding agents, as such agents (antagonists) are useful as male contraceptive agents. Similarly, the protein is believed to be useful in the treatment and/or diagnosis of testicular cancer. The testes are also a site of active gene expression of transcripts that may be expressed, particularly at low levels, in other tissues of the body. Therefore, this gene product may be expressed in other specific tissues or organs where it may play related functional roles in other processes, such as hematopoiesis, inflammation, bone formation, and kidney function, to name a few possible target indications. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:47 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 528 of SEQ ID NO:47, b is an integer of 15 to 542, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:47, and where b is greater than or equal to a + 14.

5

10

15

20

25

FEATURES OF PROTEIN ENCODED BY GENE NO: 38

5

10

15

20

25

30

It has been discovered that this gene is expressed primarily in apoptotic Tcells, and to a lesser extent in brain tissue.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: immune and neurological disorders. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and nervous systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., immune, neural, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 145 as residues: Glu-33 to Tyr-42.

The tissue distribution in apoptotic T-cells suggests potential roles in the treatment and/or diagnosis of immune disorders including of immune and autoimmune diseases, such as lupus, transplant rejection, allergic reactions, arthritis, asthma, immunodeficiency diseases, leukemia, and AIDS. Alternatively, expression in brain tissue suggests a role in the treatment and/or diagnosis of neurodegenerative disease states and behavioural disorders such as Alzheimer's Disease, Parkinson's Disease, Huntinton's Disease, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder and panic disorder. Furthermore, the tissue distribution in apoptotic T-cells indicates that the translation product of this gene may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:48 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1481 of SEQ ID NO:48, b is an integer of 15 to 1495, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:48, and where b is greater than or equal to a + 14.

10

15

20

25

30

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 39

The translation product of this gene shares sequence homology with phosphomannomutase, which is thought to be important in mannose matabolism.

It has been discovered that this gene is expressed primarily in meningioma and testis tissues.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: meningioma related diseases. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., neural, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 146 as residues: Ser-33 to Lys-43.

The tissue distribution in meningioma, and the homology to phosphomannomutase, suggests that the protein product of this clone is useful for the

diagnosis and/or intervention of meningioma related diseases. For example, the gene product can be used for preventing microbial infection of the meninges, for imaging conjugates, or as a secretory factor as a endocrine with systemic, central or peripheral nerve functions. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:49 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 804 of SEQ ID NO:49, b is an integer of 15 to 818, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:49, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 40

5

10

15

20

25

30

It has been discovered that this gene is expressed primarily in tonsils, osteoclastoma and retinoic acid treated teratocarcinoma cells, and to a lesser extent in macrophages, female bladder, adipose tissue, myeloid progenitor cells, prostate tissue, and number of other tissues and organs.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: tonsils and osteoclast related diseases. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and bone systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., immune, skeletal, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine,

10

15

20

25

30

synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 147 as residues: Glu-55 to Arg-61, Gln-84 to Ser-92, Ser-99 to Ser-104.

The tissue distribution in tonsils and osteoclastoma suggests that the protein product of this clone is useful for the diagnosis and/or intervention of diseases related to tonsils or osteoclasts. For example, tonsillitis, adenoids, peritonsilar abscess, neoplasms, or bone related disorders like rickets, abnormalities of bone growth and modelling, facture, osteonecrosis, and osteoporosis etc. Expression of this gene product in osteoclastoma suggests that it may play a role in the survival, proliferation, and/or growth of osteoclasts. Therefore, it may be useful in influencing bone mass in such conditions as osteoporosis.

Alternatively, the expression of this gene product in tonsils suggests a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are

74

related to SEQ ID NO:50 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1697 of SEQ ID NO:50, b is an integer of 15 to 1711, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:50, and where b is greater than or equal to a + 14.

10

15

20

25

30

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 41

It has been discovered that this gene is expressed primarily in resting T-cells.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: T-cell related disorders. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

The tissue distribution in resting T-cells suggests that the protein product of this clone is useful for the diagnosis and/or intervention of T-cell related disorders, such as infection, inflammation, allergy, tissue/organ transplantation, immune deficiency etc. Furthermore, the expression of this gene product in T cells also strongly suggests a role for this protein in immune function and immune surveillance. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:51 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 735 of SEQ ID NO:51, b is an integer of 15 to 749, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:51, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 42

5

10

15

The translation product of this gene shares weak sequence homology with Human metastasis suppressor KiSS-1 fragment, which is thought to be important in the diagnosis, prevention, staging and/or treatment of cancers, such as melanoma (See Accession No. W15789).

In specific embodiments, polypeptides of the invention comprise the following amino acid sequences: GQGPAGRWVRRLPCSRRAGGERGPHWGVWAGPOM 20 SCGLXFGP (SEQ ID NO:259), WRTQGPMVLLWVVTCPATMLTEPONPHLIGF VAYSGPSHTTQPHKYWLLLDGQADPAAAEGPVKRKAASVVWWPOALRHLS LLVHCWEESYEMNIGCQSLWAGGLASSGNGWDLGVAFRRDTCMSSSSLHW KEFKYAPGSLHYFALSFVLILTEICLVSSGMGFPQEGKHFSVLGSPDCSLWGR DEHVPREFA (SEQ I D NO:260), WRTQGPMVLLWVVTCPATMLTEPQNPHLIGFVAY SGPSHTTQ (SEQ ID 25 NO:261), PHKYWLLLDGQADPAAAEGPVKRKAASVVWW PQALRHLSLL (SEQ ID NO:262), VHCWEESYEMNIGCOSLWAGGLASSGNGW DLGVAFRRDTCM (SEQ I D NO:263), SSSSLHWKEFKYAPGSLHYFALSFVLILT EICLVSSGMGFPQEG (SEQ ID NO:264), and/or KHFSVLGSPDCSLWGRDEHV PREFA (SEO ID NO:265). 30 Polynucleotides encoding these polypeptides are also encompassed by the invention.

76

The gene encoding the disclosed cDNA is thought to reside on chromosome 1. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 1.

It has been discovered that this gene is expressed primarily in tonsils, osteoclastoma and teratocarcinoma tissues, and to a lesser extent in female bladder, adipose tissue, myeloid progenitor, prostate tissue, and number of other tissues.

5

10

15

20

25

30

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: diseases related to tonsils and osteoclasts. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and bone system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., immune, skeletal, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

The tissue distribution in tonsils and osteoclastoma tissues suggests that the protein product of this clone is useful for the diagnosis and/or treatment of diseases related to tonsils and osteoclasts. For example, tonsillitis, adenoids, peritonsilar abscess, neoplasms, or abnormal growth and modelling of the bone, osteonecrosis, osteoporosis, osteodystrophy, osteoclastoma etc. Expression of this gene product in osteoclastoma suggests that it may play a role in the survival, proliferation, and/or growth of osteoclasts. Therefore, it may be useful in influencing bone mass in such conditions as osteoporosis.

Moreover, the expression of this gene product in tonsils suggests a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen

10

15

20

30

presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:52 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1077 of SEQ ID NO:52, b is an integer of 15 to 1091, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:52, and where b is greater than or equal to a + 14.

25 FEATURES OF PROTEIN ENCODED BY GENE NO: 43

The translation product of this gene shares sequence homology with the Drosophila gene "maleless", which is one of four known regulatory loci required for increased transcription (dosage compensation) of X-linked genes (See Genbank Accession No.: gil157906).

It has been discovered that this gene is expressed primarily in normal prostate tissue, testes tissue, whole 6-week old embryonic tissue, human colon carcinoma

10

15

20

25

30

(HCC) cell line, and cerebellum tissue, and to a lesser extent in primary breast cancer, activated T-cells, and many other tissues.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: diseases of the prostate or colon, or male reproductive disorders. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the prostate or colon carcinoma, and male reproductive disorders, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., colon, prostate, reproductive, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 150 as residues: Val-39 to Ala-45.

The tissue distribution in colon and prostate tissues suggests that the protein product of this clone is useful for the diagnosis and/or treatment of prostate disorders such as prostatitis, prostatic hyperplasia, prostate cancers, or human colon carcinoma, as well as cancers of other tissues where expression has been observed. Alternatively, the tissue distribution in testes tissue, in conjunction with the homology to the Drosophila maleless gene, suggests that the translation product of this gene is useful for the detection and/or treatment of disorders involving the testes or the transcription of X-linked genes. Furthermore, the tissue distribution indicates that the protein product of this clone is useful for the treatment and diagnosis of conditions concerning proper testicular function (e.g. endocrine function, sperm maturation), as well as cancer. Therefore, this gene product is useful in the treatment of male infertility and/or impotence.

This gene product is also useful in assays designed to identify binding agents, as such agents (antagonists) are useful as male contraceptive agents. Similarly, the

WO 99/47540

protein is believed to be useful in the treatment and/or diagnosis of testicular cancer. The testes are also a site of active gene expression of transcripts that may be expressed, particularly at low levels, in other tissues of the body. Therefore, this gene product may be expressed in other specific tissues or organs where it may play related functional roles in other processes, such as hematopoiesis, inflammation, bone formation, and kidney function, to name a few possible target indications. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:53 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2240 of SEQ ID NO:53, b is an integer of 15 to 2254, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:53, and where b is greater than or equal to a + 14.

20

25

30

15

5

10

FEATURES OF PROTEIN ENCODED BY GENE NO: 44

The translation product of this gene shares weak sequence homology with Eimeria antigen Eam45 M3, which is thought to be important in uses as a vaccine for protecting chickens against coccidiosis.

It has been discovered that this gene is expressed primarily in adrenal gland tissue, and to a lesser extent in activated T-cells.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: adrenal cortical insufficiency, adrenal cortical hyperfunction, neoplasia. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for

80

differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., endocrine, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

The tissue distribution in adrenal gland tissue suggests that the protein product of this clone is useful for the diagnosis and/or intervention of disorders caused by adrenal gland abnormalities, such as adrenal cortical insufficiency, adrenal cortical hyperfunction, and neoplasia. More generally, the tissue distribution suggests that the protein product of this clone is useful for the detection, treatment, and/or prevention of various endocrine disorders and cancers, particularly Addison's disease, Cushing's Syndrome, and disorders and/or cancers of the pancrease (e.g. diabetes mellitus), adrenal cortex, ovaries, pituitary (e.g., hyper-, hypopituitarism), thyroid (e.g. hyper-, hypothyroidism), parathyroid (e.g. hyper-, hypoparathyroidism), hypothallamus, and testes. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:54 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 472 of SEQ ID NO:54, b is an integer of 15 to 486, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:54, and where b is greater than or equal to a + 14.

30

10

15

20

25

10

15

20

The translation product of this gene shares sequence homology with neural thread protein, tumor necrosis factor related gene product, human alpha-1C2 adrenalin receptor, which is thought to be important for diagnosing the presence of Alzheimer's disease, neuroectodermal tumours and a malignant astrocytoma, or diagnosis of hepatocellular carcinomas and preneoplastic or pathological conditions of the liver, and tumor immunity.

It has been discovered that this gene is expressed primarily in activated T-cells and endothelial cells.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: Alzheimer's disease, neuroectodermal tumours and a malignant astrocytoma, hepatocellular carcinomas and tumors of various origins. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system and endothelial cells, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., immune, endothelial, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 152 as residues: Arg-38 to Arg-47.

The tissue distribution in immune and endothelial tissues, and the homology to

neural thread protein, tumor necrosis factor related gene product, human alpha-1C2

adrenalin receptor, or Smaller hepatocellular oncoprotein (hhcm) gene product

suggests that the protein product of this clone is useful for the diagnosis and/or

treatment of tumors of various origins, including neuroectodermal tumours and a

malignant astrocytoma, hepatocellular carcinomas, as well as syndromes inflicted by

these cancers. Protein, as well as, antibodies directed against the protein may show

utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

15

20

25

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:55 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1256 of SEQ ID NO:55, b is an integer of 15 to 1270, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:55, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 46

It has been discovered that this gene is expressed primarily in tumor tissues such as hepatocellular tumor, hemangiopericytoma, chronic lymphocytic leukemia, and activated T-cells.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: tumors of various origins. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hepatocellular tumor, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., liver, immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

The tissue distribution in hepatocellular tumors suggests that the protein

30 product of this clone is useful for the diagnosis and/or targeting of hepatocellular carcinomas, preneoplastic or pathological conditions of the liver, Alzheimer's disease,

neuroectodermal tumours and malignant astrocytoma. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:56 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2045 of SEQ ID NO:56, b is an integer of 15 to 2059, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:56, and where b is greater than or equal to a + 14.

15

20

25

30

10

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 47

It has been discovered that this gene is expressed primarily in glioblastoma, ulcerative colitis, and hemangiopericytoma.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: glioblastoma, hemangiopericytoma and their inflicted disorders. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain tissues, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., neural, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

84

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 154 as residues: Pro-31 to Ala-37.

The tissue distribution suggests that the protein product of this clone would be useful for the diagnosis, targeting and/or treatment of tumors in the brain, such as glioblastoma and hemangiopericytoma. Additionally, the gene products can be useful agent for the diagnosis and treatment of ulcerative colitis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:57 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 854 of SEQ ID NO:57, b is an integer of 15 to 868, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:57, and where b is greater than or equal to a + 14.

20

25

30

5

10

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 48

It has been discovered that this gene is expressed primarily in bone marrow.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: immunodeficiency, tumor necrosis, infection, lymphomas, auto-immunities, cancer, inflammation, anemias (leukemia) and other hematopoeitic disorders. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types

10

15

20

25

30

PCT/US99/05804

(e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 155 as residues: Thr-47 to Val-53.

The tissue distribution in bone marrow suggests that the protein product of this clone is useful for the diagnosis and/or treatment of immune disorders including: leukemias, lymphomas, auto-immunities, immunodeficiencies (e.g. AIDS), immunosupressive conditions (transplantation) and hematopoeitic disorders. In addition this gene product may be applicable in conditions of general microbial infection, inflammation or cancer. Furthermore, the tissue distribution in bone marrow suggests that the protein product of this clone is useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia.

The uses include bone marrow cell ex vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:58 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the

general formula of a-b, where a is any integer between 1 to 972 of SEQ ID NO:58, b is an integer of 15 to 986, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:58, and where b is greater than or equal to a + 14.

5

10

15

20

25

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 49

It has been discovered that this gene is expressed primarily in bone marrow.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: immunodeficiency, tumor necrosis, infection, lymphomas, auto-immunities, cancer, inflammation, anemias (leukemia) and other hematopoeitic disorders. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 156 as residues: Leu-40 to Cys-47.

The bone marrow tissue distribution suggests that the protein product of this clone would be useful for the diagnosis and treatment of immune disorders including: leukemias, lymphomas, auto-immunities, immunodeficiencies (e.g. AIDS), immunosupressive conditions (transplantation) and hematopoeitic disorders. In addition this gene product may be applicable in conditions of general microbial infection, inflammation or cancer. Furthermore, the tissue distribution in bone marrow suggests that the protein product of this clone is useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia.

The uses include bone marrow cell ex vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

5

10

15

20

25

30

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:59 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 681 of SEQ ID NO:59, b is an integer of 15 to 695, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:59, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 50

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: IAQGTVPLTKRGVQSSGPDYPEGTLTPLPRG (SEQ ID NO:266 and 267). Polynucleotides encoding these polypeptides are also encompassed by the invention.

It has been discovered that this gene is expressed primarily in dendritic cells.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: immune disorders and related conditions such as leukemias, lymphomas, inflammation, hematopoeitic disfunction,

15

20

25

30

arthritis and asthma. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of dendritic cells. For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., dendritic cells, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 157 as residues: Ser-25 to Phe-31, Lys-55 to Arg-61.

The tissue distribution in dendritic cells suggests that the protein product of this clone is useful for the diagnosis and/or treatment of immune disorders including: leukemias, lymphomas, auto-immunities, immunodeficiencies (e.g. AIDS), immunosupressive conditions (transplantation) and hematopoeitic disorders. In addition this gene product may be applicable in conditions of general microbial infection, inflammation or cancer.

Moreover, the expression of this gene product in dendritic cells also strongly suggests a role for this protein in immune function and immune surveillance. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:60 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 300 of SEQ ID NO:60, b is an integer of 15 to 314, where both a and b correspond to the positions of

20

25

30

nucleotide residues shown in SEQ ID NO:60, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 51

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: DCLYLALSFPWHCHCHHHPPSGSLLYPF (SEQ ID NO:268). Polynucleotides encoding these polypeptides are also encompassed by the invention. The translation product of this gene shares sequence homology with a C. elegans protein of unknown function (See Genbank Accession No.: gil1947142 (AF000264)).

It has been discovered that this gene is expressed primarily in healing abdominal wound tissue.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: tissue necrosis, wound healing, ulceration, neoplasms or cancer. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of injured tissue, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., vascular, endothelial, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 158 as residues: Pro-34 to Tyr-43, Gln-73 to Cys-86, Pro-98 to Leu-103.

The tissue distribution in healing abdominal wound tissue suggests that the protein product of this clone is useful for the treatment and/or diagnosis of conditions involving tissue repair and wound healing. Tissue repair may be indicated in cases of injury to the skin or internal organs, ulceration, cellular necrosis or other conditions involving healing of both diseased or non-diseased, traumatized tissue. In addition,

90

because of the implications of tissue regeneration, remoldeling and growth regulation, the protein product of this gene may have indications in the diagnosis and treatment of neoplasms and cancer.

5

10

15

20

25

30

More generally, the tissue distribution in endothelial tissue indicates that the protein product of this gene is useful for the diagnosis and treatment of conditions and pathologies of the cardiovascular system, such as heart disease, restenosis, atherosclerosis, stoke, angina, thrombosis, and wound healing. Likewise, the tissue distribution further suggests that the protein product of this clone is useful for the treatment, diagnosis, and/or prevention of various skin disorders including congenital disorders (i.e. nevi, moles, freckles, Mongolian spots, hemangiomas, port-wine syndrome), integumentary tumors (i.e. keratoses, Bowen's disease, basal cell carcinoma, squamous cell carcinoma, malignant melanoma, Paget's disease, mycosis fungoides, and Kaposi's sarcoma), injuries and inflammation of the skin (i.e.wounds, rashes, prickly heat disorder, psoriasis, dermatitis), atherosclerosis, uticaria, eczema, photosensitivity, autoimmune disorders (i.e. lupus erythematosus, vitiligo, dermatomyositis, morphea, scleroderma, pemphigoid, and pemphigus), keloids, striae, erythema, petechiae, purpura, and xanthelasma. In addition, such disorders may predispose increased susceptibility to viral and bacterial infections of the skin (i.e. cold sores, warts, chickenpox, molluscum contagiosum, herpes zoster, boils, cellulitis, erysipelas, impetigo, tinea, althletes foot, and ringworm). Moreover, the protein product of this clone may also be useful for the treatment or diagnosis of various connective tissue disorders such as arthritis, trauma, tendonitis, chrondomalacia and inflammation, autoimmune disorders such as rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism, spinal deformation, and specific joint abnormalities as well as chondrodysplasias (i.e. spondyloepiphyseal dysplasia congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are

related to SEQ ID NO:61 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 720 of SEQ ID NO:61, b is an integer of 15 to 734, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:61, and where b is greater than or equal to a + 14.

10

15

20

25

30

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 52

The translation product of this gene shares sequence homology with FAR-17A, which is an androgen induced protein, absent in castrated hamsters (See Genbank Accession No.: gil191315), as well as a male hormone-dependent gene product (See GenSeq Accession No.: R10612). The gene encoding the disclosed cDNA is thought to reside on chromosome 6. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 6.

In specific embodiments, polypeptides of the invention comprise the following amino acid sequences: ASLPPSRSRPLANMALVPCQVLRMAILLSYCSILCNYKA IEMPSHQTYGGSWKFLTFIDLVIQAVFFGICVLTDLSSLLTRGSGNQEQERQLK KLISLRDW (SEQ ID NO:269). Polynucleotides encoding these polypeptides are also encompassed by the invention.

It has been discovered that this gene is expressed primarily in fetal liver and spleen tissue, and to a lesser extent in a variety of other fetal tissues and brain tissues.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: immune disorders including leukemias, lymphomas; reproductive and endocrine disorders, including testicular cancer; and liver disorders (e.g. hepatoblastoma, metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells). Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide

WO 99/47540

10

15

20

25

30

PCT/US99/05804

immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and reproductive systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., immune, reproductive, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

92

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 159 as residues: Thr-59 to Gly-70, Tyr-132 to Glu-150.

The tissue distribution and homology to FAR-17A suggests that the protein product of this clone is useful for the treatment and/or diagnosis of androgen related conditions and disorders. Male reproductive and endocrine disorders would be potential area of application (e.g. endocrine function, sperm maturation). It may also prove to be valuable in the diagnosis and treatment of testicular cancer.

More generally, the protein product of this clone may be useful for the treatment and/or diagnosis of conditions concerning proper testicular function (e.g. endocrine function, sperm maturation), as well as cancer. Therefore, this gene product is useful in the treatment of male infertility and/or impotence. This gene product is also useful in assays designed to identify binding agents, as such agents (antagonists) are useful as male contraceptive agents. Similarly, the protein is believed to be useful in the treatment and/or diagnosis of testicular cancer. The testes are also a site of active gene expression of transcripts that may be expressed, particularly at low levels, in other tissues of the body. Therefore, this gene product may be expressed in other specific tissues or organs where it may play related functional roles in other processes, such as hematopoiesis, inflammation, bone formation, and kidney function, to name a few possible target indications. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are

related to SEQ ID NO:62 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1396 of SEQ ID NO:62, b is an integer of 15 to 1410, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:62, and where b is greater than or equal to a + 14.

10

15

20

25

30

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 53

Contact of cells with supernatant expressing the product of this gene has been shown to increase the permeability of the plasma membrane of THP-1 to calcium. Thus it is likely that the product of this gene is involved in a signal transduction pathway that is initiated when the product binds a receptor on the surface of the plasma membrane of monocytes, and to a lesser extent, in immune or hematopoietic cells and tissues. Thus, polynucleotides and polypeptides have uses which include, but are not limited to, activating monocytes.

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: MSRSSRISGLSCPWLL (SEQ ID NO:270). Polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is believed to reside on chromosome 1. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 1.

It has been discovered that this gene is expressed primarily in T-cells.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: immune and hematopoietic diseases and/or disorders. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells,

94

particularly of the immune and haemopoietic systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

5

10

15

20

25

30

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 160 as residues: Pro-42 to Cys-50, Leu-61 to Ala-66.

The tissue distribution in T-cells, combined with the detected calcium flux activity in monocytes suggests that the protein product of this clone would be useful for the treatment and diagnosis of immune and hematopoietic disorders. Morever, the expression of this gene product suggests a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosis, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues.

Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as,

PCT/US99/05804

antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:63 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1217 of SEQ ID NO:63, b is an integer of 15 to 1231, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:63, and where b is greater than or equal to a + 14.

15 FEATURES OF PROTEIN ENCODED BY GENE NO: 54

5

10

20

25

30

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: DHWPAGFLPPAPGLKFPVALEVFRKVLPAVCPTDCSGS AGKERNS (SEQ ID NO:271). Polynucleotides encoding these polypeptides are also encompassed by the invention.

It has been discovered that this gene is expressed primarily in liver.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: metabolic diseases and liver conditions. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the metabolic system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., hepatic, liver, metabolic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, bile, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder,

10

15

20

30

PCT/US99/05804

relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 161 as residues: Ser-31 to Gln-41.

The tissue distribution in liver suggests that the protein product of this clone would be useful for treatment and diagnosis of disorders of the metabolic system and liver disorders. Morever, the protein product of this clone is useful for the detection and treatment of liver disorders and cancers (e.g. hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:64 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 598 of SEQ ID NO:64, b is an integer of 15 to 612, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:64, and where b is greater than or equal to a + 14.

25 FEATURES OF PROTEIN ENCODED BY GENE NO: 55

When tested against PC12 cell lines, supernatants removed from cells containing this gene activated the EGR1 (early growth response gene 1) promoter element. Thus, it is likely that this gene activates sensory neuron cells, and to a lesser extent in other neural cells and tissues, through the EGR1 signal transduction pathway. EGR1 is a separate signal transduction pathway from Jak-STAT, genes

containing the EGR1 promoter are induced in various tissues and cell types upon activation, leading the cells to undergo differentiation and proliferation.

It has been discovered that this gene is expressed primarily in T-cells and monocytes, and to a lesser extent in cancerous tissues, including cancerous colon tissue and placenta.

5

10

15

20

25

30

j

2

25

30

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: immune and haemopoietic disorders and cancer such as colon cancer, but also such cancers as breast cancer, cardiac tumors, pancreatic cancer, melanoma, retinoblastoma, glioblastoma, lung cancer, intestinal cancer, testicular cancer, stomach cancer, neuroblastoma, myxoma, myoma, lymphoma, endothelioma, osteoblastoma, osteoclastoma, adenoma, and the like. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and haemopoietic systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 162 as residues: Glu-63 to Trp-72.

The tissue distribution in T-cells and monocytes, combined with the detected EGR1 biological activity suggests that the protein product of this clone would be useful for treatment and diagnosis of disorders of the immune and haemopoietic systems and colon and other cancers. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an

agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosis, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues.

5

10

15

20

25

30

98

Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Expression cellular sources marked by proliferating cells suggests this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA).

Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and would be useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:65 and may have been publicly available prior to conception of

the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2256 of SEQ ID NO:65, b is an integer of 15 to 2270, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:65, and where b is greater than or equal to a + 14.

99

10 FEATURES OF PROTEIN ENCODED BY GENE NO: 56

5

15

20

25

30

The translation product of this gene has homology with several human keratin genes at the nucleotide level (see, for example, Troyanovsky, et al., Eur. J. Cell Biol. 59:127-137 (1992) which is hereby incorporated by reference herein). Based on the sequence similarity, the translation product of this clone is expected to share biological activities with keratin and growth factor proteins. Such activities are known in the art, and some of which are described elsewhere herein.

It has been discovered that this gene is expressed primarily in neutrophils.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: immune and haemopoietic disorders. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and haemopoietic system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

The tissue distribution in neutrophils suggests that the protein product of this clone would be useful for treatment and diagnosis of disorders of the immune and

haemopoietic system. Furthermore, sequence homology of the polynucleotides and polypeptides of the present invention with a number of human cytokeratin molecules, such as CK-8, CK-15, and CK-17, indicate that molecules of the present invention can be used diagnostically as markers of basal cell differentiation in complex epithelia and therefore indicative of a certain type of epithelial stem cells, as well as markers of the differentiation of other cell types such as neutrophils or other immune cells. Molecules of the present invention, or agonists or antagonists thereof, can also be used therapeutically to treat differentiation disorders of epithelial, neutrophil or other immune cell differentiation or activation. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:66 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1269 of SEQ ID NO:66, b is an integer of 15 to 1283, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:66, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 57

10

15

20

25

30

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: EEIATSIEPIRDFLAIVFFASIGLHVFPTFVAYELTVLVF LTLSVVV (SEQ ID NO:272). Polynucleotides encoding these polypeptides are also encompassed by the invention.

It has been discovered that this gene is expressed primarily in synovium, placenta, and stromal cells, and to a lesser extent in several other tissues and organs,

WO 99/47540

5

10

15

20

including, among others, bone marrow, palate, pituitary gland, and in tissue derived from osteosarcoma and chondrosarcoma.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: developmental disorders, as well as disorders of the musculoskeletal and haematopoietic systems, and cancers including especially osteosarcoma and chondrosarcoma, but also other cancers including breast cancer, colon cancer, cardiac tumors, pancreatic cancer, melanoma, retinoblastoma, glioblastoma, lung cancer, intestinal cancer, testicular cancer, stomach cancer, neuroblastoma, myxoma, myoma, lymphoma, endothelioma, osteoblastoma, osteoclastoma, adenoma, and the like. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the haemopoietic and musculoskeletal systems, as well as developmental disorders, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., synovium, placenta, stromal, immune, hematopoietic, skeletal, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 164 as residues: Pro-81 to Ser-88.

The tissue distribution in placenta suggests that the protein product of this

clone would be useful for treatment and diagnosis of developmental disorders.

Polynucleotides and polypeptides of the present invention can be used diagnostically and therapeutically to detect and treat many cancers, particularly osteosarcoma and chondrosarcoma. In addition, the expression of this gene product in synovium would suggest a role in the detection and treatment of disorders and conditions affecting the

skeletal system, in particular osteoporosis, bone cancer, as well as, disorders afflicting connective tissues (e.g. arthritis, trauma, tendonitis, chrondomalacia and

10

15

20

25

30

inflammation), such as in the diagnosis or treatment of various autoimmune disorders such as rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism, spinal deformation, and specific joint abnormalities as well as chondrodysplasias (i.e. spondyloepiphyseal dysplasia congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid).

Moreover, the protein is useful in the detection, treatment, and/or prevention of a variety of vascular disorders and condtions, which include, but are not limited to miscrovascular disease, vascular leak syndrome, aneurysm, stroke, embolism, thrombosis, coronary artery disease, arteriosclerosis, and/or atherosclerosis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:67 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1249 of SEQ ID NO:67, b is an integer of 15 to 1263, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:67, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 58

Contact of cells with supernatant expressing the product of this gene has been shown to increase the permeability of the plasma membrane of renal messiaglia cells to calcium. Thus it is likely that the product of this gene is involved in a signal transduction pathway that is initiated when the product binds a receptor on the surface of the plasma membrane of renal and developing cells and tissues. Thus, polynucleotides and polypeptides have uses which include, but are not limited to, activating renal and developing cells and tissues.

103

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: YCNLQCR (SEQ ID NO:273). Polynucleotides encoding these polypeptides are also encompassed by the invention.

It has been discovered that this gene is expressed primarily in the whole developing embryo, as well as in ovarian cancer and placenta.

5

10

15

20

25

30

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: developmental or reproductive diseases and/or disorders, in addition to the following and ovarian cancer, as well as other cancers including breast cancer, colon cancer, cardiac tumors, pancreatic cancer, melanoma, retinoblastoma, glioblastoma, lung cancer, intestinal cancer, testicular cancer, stomach cancer, neuroblastoma, myxoma, myoma, lymphoma, endothelioma, osteoblastoma, osteoclastoma, osteosarcoma, chondrosarcoma, adenoma, and the like. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the developing and fetal system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., developmental, reproductive, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

The tissue distribution in embryonic and ovarian tissue, combined with the detected calcium flux activity, suggests that the protein product of this clone would be useful for tretment and diagnosis of developmental disorders as well as ovarian and other cancers. Expression within embryonic tissue and other cellular sources marked by proliferating cells suggests this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of

some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA).

Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and would be useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases.

Alternatively, the protein is useful in the detection, treatment, and/or prevention of vascular conditions, which include, but are not limited to, microvascular disease, vascular leak syndrome, aneurysm, stroke, atherosclerosis, arteriosclerosis, or embolism. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:68 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1603 of SEQ ID NO:68, b is an integer of 15 to 1617, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:68, and where b is greater than or equal to a + 14.

25

30

5

10

15

20

FEATURES OF PROTEIN ENCODED BY GENE NO: 59

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: SALIGNPKGCFGCFSPVVLREWSVESWKSLRPFQAICK LKTNFR (SEQ ID NO:274). Polynucleotides encoding these polypeptides are also encompassed by the invention.

15

20

25

30

It has been discovered that this gene is expressed primarily in hypothalamus and anergic T cells.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: neurological and inflammatory defects, diseases, and/or disorders. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous and immune systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues (e.g., neural, immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 166 as residues: His-33 to Trp-38.

The tissue distribution in hypothalamus and T-cells suggests that the protein product of this clone would be useful for study and treatment of immune and nervous system disorders. The protein product of this clone is useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions which include, but are not limited to Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated

Expression of this gene product in regions of the brain suggests it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal

10

15

20

25

30

differentiation or survival. Morever, the expression of this gene product suggests a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosis, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:69 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1375 of SEQ ID NO:69, b

is an integer of 15 to 1389, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:69, and where b is greater than or equal to a + 14.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 60

10

15

20

25

30

The translation product of this gene shares nucleotide sequence homology with the human PKD1 gene which is thought to be important in polycystic kidney disease.

This gene is expressed widely with a predominant expression exhibited in liver, pediatric kidney, and in the whole 8 week old developing human embryo.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: cancer, growth, renal, and metabolic defects, diseases, and/or disorders. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine, digestive and immune systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., renal, metabolic, hepatic, developmental, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid, bile, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

The tissue distribution in pediatric kidney suggests that the protein product of this clone would be useful for study and treatment of renal and general neoplasias and growth and development disorders. The protein product of this clone could be used in the treatment and/or detection of kidney diseases including renal failure, nephritus, renal tubular acidosis, proteinuria, pyuria, edema, pyelonephritis, hydronephritis, nephrotic syndrome, crush syndrome, glomerulonephritis, hematuria, renal colic and kidney stones, in addition to Wilm's Tumor Disease, and congenital kidney abnormalities such as horseshoe kidney, polycystic kidney, and Falconi's syndrome.

10

15

20

25

30

Moreover, the expression within embryonic tissue suggests this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders, particularly of the liver and other organs. Similarly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and would be useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:70 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1882 of SEQ ID NO:70, b is an integer of 15 to 1896, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:70, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 61

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: HEAALRGP (SEQ ID NO:275). Polynucleotides encoding these polypeptides are also encompassed by the invention.

5

10

15

20

25

30

109

PCT/US99/05804

It has been discovered that this gene is expressed primarily in human striatum depression.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: stroke, in addition to other, neurologically-related diseases and/or defects. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be detected in certain tissues (e.g., neural, musculoskeletal, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 168 as residues: Glu-50 to Glu-61.

The tissue distribution in human striatum depression suggests that the protein product of this clone would be useful for study and treatment of central nervous system orders, such as seizures and other neurological conditions. The protein product of this clone is useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions which include, but are not limited to Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated

Expression of this gene product in regions of the brain suggests it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal

differentiation or survival. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:71 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 294 of SEQ ID NO:71, b is an integer of 15 to 308, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:71, and where b is greater than or equal to a + 14.

15

20

25

30

10

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 62

This clone has homology to a cystine rich granulin peptide(s) from leucocyte(s) which has been termed Granulin E. Granulins inhibit keratinocytes and is useful topically for wound healing. The gene encoding the disclosed cDNA is believed to reside on chromosome 3. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 3.

It has been discovered that this gene is expressed primarily in infant brain.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: neurological, developmental, and growth defects. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the fetus and the nervous system, expression of this gene at significantly higher or lower levels may be detected in certain tissues (e.g., neural, developmental, growth, and cancerous and wounded tissues) or bodily fluids (e.g.,

5

10

15

20

25

30

.

PCT/US99/05804

lymph, amniotic fluid, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder. Based on the strong conservation of cysteine residues, the polypeptide of the present invention can be used to inhibit keratinocytes and promote wound healing.

111

The tissue distribution in infant brain suggests that the protein product of this clone would be useful for study and treatment of nervous system, neurodegenerative and developmental disorders. The protein product of this clone is useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions which include, but are not limited to Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated

Expression of this gene product in regions of the brain suggests it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. The homology to granulin proteins suggest the protein product of this clone is useful for the treatment, diagnosis, and/or prevention of various skin disorders including congenital disorders (i.e. nevi, moles, freckles, Mongolian spots, hemangiomas, port-wine syndrome), integumentary tumors (i.e. keratoses, Bowen's disease, basal cell carcinoma, squamous cell carcinoma, malignant melanoma, Paget's disease, mycosis fungoides, and Kaposi's sarcoma), injuries and inflammation of the skin (i.e. wounds, rashes, prickly heat disorder, psoriasis, dermatitis), atherosclerosis, uticaria, eczema, photosensitivity, autoimmune disorders (i.e. lupus erythematosus, vitiligo, dermatomyositis, morphea, scleroderma, pemphigoid, and pemphigus),

WO 99/47540 PCT/US99/05804

112

keloids, striae, erythema, petechiae, purpura, and xanthelasma. In addition, such disorders may predispose increased susceptibility to viral and bacterial infections of the skin (i.e. cold sores, warts, chickenpox, molluscum contagiosum, herpes zoster, boils, cellulitis, erysipelas, impetigo, tinea, althletes foot, and ringworm). Moreover, the protein product of this clone may also be useful for the treatment or diagnosis of various connective tissue disorders such as arthritis, trauma, tendonitis, chrondomalacia and inflammation, autoimmune disorders such as rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism, spinal deformation, and specific joint abnormalities as well as chondrodysplasias (i.e. spondyloepiphyseal dysplasia congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:72 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1674 of SEQ ID NO:72, b is an integer of 15 to 1688, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:72, and where b is greater than or equal to a + 14.

25

30

20

10

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 63

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: SNAAGNVVRAFLYINHLKL GCKVGLA (SEQ ID NO:276). Polynucleotides encoding these polypeptides are also encompassed by the invention.

It has been discovered that this gene is expressed primarily in prostate cancer and dendritic cells.

10

15

20

25

30

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: reproductive, immune, and hematopoietic diseases, defects and/or disorders. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine and immune systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., reproductive, immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, seminal fluid, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 170 as residues: Trp-47 to Thr-54.

The tissue distribution in prostate cells and tissues indicates that the protein products of this clone are useful for study, diagnosis and treatment of neoplasias, esp. of the prostate, and hormonal and metabolic disorders. Moreover, the protein product of this clone is useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include bone marrow cell ex- vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

114

PCT/US99/05804

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:73 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1124 of SEQ ID NO:73, b is an integer of 15 to 1138, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:73, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 64

5

10

15

20

25

30

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: NWAVLNMLLSKGKITIFLGPLECGS (SEQ ID NO:277).

Polynucleotides encoding these polypeptides are also encompassed by the invention.

It has been discovered that this gene is expressed primarily in B cell lymphoma.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: immune and hematopoietic diseases, disorders, and/or defects, particularly cancers. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hemopoietic and immune systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

15

20

25

The tissue distribution in B cell lymphoma suggests that the protein product of this clone would be useful for study and treatment of blood and immune disorders and neoplasias, esp. of the lymphatic system. The protein product of this clone is useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include bone marrow cell ex- vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:74 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 763 of SEQ ID NO:74, b is an integer of 15 to 777, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:74, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 65

It has been discovered that this gene is expressed primarily in B cell lymphoma.

WO 99/47540 PCT/US99/05804

5

10

15

20

25

116

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: immune and hematopoietic diseases, disorders, and/or defects, particularly cancer. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hemopoietic and immune systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

The tissue distribution in B cell lymphoma suggests that the protein product of this clone would be useful for study and treatment of neplasias, esp. of lymphatic organs, and immune disorders. The protein product of this clone is useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include bone marrow cell ex- vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly

available and accessible through sequence databases. Some of these sequences are
related to SEQ ID NO:75 and may have been publicly available prior to conception of

PCT/US99/05804

the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1046 of SEQ ID NO:75, b is an integer of 15 to 1060, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:75, and where b is greater than or equal to a + 14.

10 FEATURES OF PROTEIN ENCODED BY GENE NO: 66

5

15

20

The translation product of this gene shares sequence homology with a rat protein phosphatase, in addition to, a human heterogeneous nuclear ribonucleoprotein R (See Genbank Accession No.gil2697103 (AF000364)). When tested against PC12 cell lines, supernatants removed from cells containing this gene activated the EGR1 (early growth response gene 1) promoter element. Thus, it is likely that this gene activates sensory neuron cells through the EGR1 signal transduction pathway. EGR1 is a separate signal transduction pathway from Jak-STAT, genes containing the EGR1 promoter are induced in various tissues and cell types upon activation, leading the cells to undergo differentiation and proliferation. This gene also showed activity in sensory neurons using the EGR assay described in the Example section.

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: PSHQTRKGKSAKLLDRPPEALRMKIITTTLLLACHLQLEV G V V V G G E V D (S E Q I D N 0:278), FQASSANNQQNWGSQPIAQQPLQQGGDYSG

25 NYGYNNDNQEFYQDTYGQQWK (SEQ ID NO:279), WXPLLXTSGSPGLXGFG TRMNGKEIEGEEIEIVLAKPPDKKRKERQAARQASRSTAYEDYYYHPPPRMPP PIRGRGRGGGGGGGGYGYPPDYYGYEDYYDDYYGYDYHDYRGGYEDPYYGYD DGYAVRGRGGGRGAPPPPRGRGAPPPRGRAGYSQRGAPLGPPRGSRGG RGGPAQQQRGRGSRGSRGNRGGNVGGKRKADGYNQPDSKRRQPTTNRTGV PNPSLSSRFSKVVTILVTMVTIMTTRNFIRILMGNSGSRQVRA (SEQ ID NO:280), RMNGKEIEGEEIEIVLAKPPDKKRKER (SEQ ID NO:281), YYHPPP

10

15

20

25

30

It has been discovered that this gene is expressed primarily in human primary breast cancer, lung, and leukocytes.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: reproductive, immune, or pulmonary diseases and/or disorders, particularly breast cancer. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive, immune and respiratory systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., reproductive, immune, pulmonary, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

The tissue distribution in breast cancer cells and tissues, in addition to immune cells, combined with the homology to a protein phosphatase suggests that the protein product of this clone would be useful for diagnosis and treatment of breast cancer and abnormalities of the lung and the immune system. Morever, the expression of this gene product suggests a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

PCT/US99/05804

5

10

15

20

25

30

Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosis, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues.

Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. The protein is useful in modulating the immune response to aberrant cells and cell types, particularly proliferative cells (e.g. protein may increase the immunogenicity of tumor antigens either directly or indirectly, or may activate apoptosis). The protein is useful in treating, detecting, and/or preventing various pulmonary disorders, which include, but are not limited to, ARDS, emphysema, and cystic fibrosis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:76 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1489 of SEQ ID NO:76, b is an integer of 15 to 1503, where both a and b correspond to the positions of

10

15

20

25

30

nucleotide residues shown in SEQ ID NO:76, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 67

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: LQIPPSSQSLGLKNADSSI (SEQ ID NO:286), GGPPESAPW LPAVLRAPVLTSRCASSDSEGPVWFCQPGSGPSSTEMSCHCILGPGSSCLCVL RGSMWTPSVPGWPQPAKETGASSCSVFSANNGSCPLPLHNHQRQASLDTGL SLEHVPGESYFYSPVG (SEQ ID NO:287), SSDSEGPVWFCQPGSGPSSTEMSC HCILGPGSSC (SEQ ID NO:288), WTPSVPGWPQPAKETGASSCSVFSANNG (SEQ ID NO:289), and/or QRQASLDTGL SLEHVPGESYF (SEQ ID NO:290). Polynucleotides encoding these polypeptides are also encompassed by the invention.

It has been discovered that this gene is expressed primarily in human B cell lymphoma.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: immune or hematopoietic diseases and/or disorders, particularly B cell lymphoma. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

The tissue distribution in B-cell lymphoma suggests that the protein product of this clone would be useful for diagnosis and treatment of immune or hematopoietic diseases and/or disorders, particularly proliferative conditions. Morever, the expression of this gene product suggests a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including

WO 99/47540 PCT/US99/05804

121

5

10

15

20

25

30

blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosis, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. The uses include bone marrow cell ex- vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:77 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence

would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 858 of SEQ ID NO:77, b is an integer of 15 to 872, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:77, and where b is greater than or equal to a + 14.

122

PCT/US99/05804

FEATURES OF PROTEIN ENCODED BY GENE NO: 68

5

10

15

20

25

30

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: SSSLVLTIRSQTLFLASFIHSTSIFCALN (SEQ ID NO:291). Polynucleotides encoding these polypeptides are also encompassed by the invention.

It has been discovered that this gene is expressed primarily in osteoarthritic cartilage.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: osteoarthritis and other bone/cartilage disorders, particularly degenerative conditions. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of these tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skelatal system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., skeletal, joint, autoimmune, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

The tissue distribution in osteoarthritic cartilage suggests that the protein product of this clone would be useful for the diagnosis, treatment, and/or prevention of osteoarthritis. Moreover, the gene product is useful in the detection and treatment of disorders and conditions affecting the skeletal system, in particular osteoporosis, bone cancer, as well as, disorders afflicting connective tissues (e.g. arthritis, trauma,

tendonitis, chrondomalacia and inflammation), such as in the diagnosis or treatment of various autoimmune disorders such as rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism, spinal deformation, and specific joint abnormalities as well as chondrodysplasias (i.e. spondyloepiphyseal dysplasia congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:78 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 559 of SEQ ID NO:78, b is an integer of 15 to 573, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:78, and where b is greater than or equal to a + 14.

20

25

30

5

10

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 69

The gene encoding the disclosed cDNA is believed to reside on chromosome 17. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 17.

It has been discovered that this gene is expressed primarily in fetal brain, pharynx carcinoma, and Hodgkin's lymphoma.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: developmental and/or proliferative diseases and disorders, particularly pharynx carcinoma, and Hodgkin's lymphoma. Similarly, polypeptides and antibodies directed to those polypeptides are useful to

WO 99/47540 PCT/US99/05804

124

provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the digestive and immune systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., developmental, proliferative cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, amniotic fluid, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

5

10

15

20

25

30

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 176 as residues: Tyr-30 to Ser-40.

The tissue distribution in pharynx carcinoma and Hodgkin's lymphoma suggests that the protein product of this clone would be useful for diagnosis and treatment of immune and proliferative conditions. Moreover, expression within fetal tissue and other cellular sources marked by proliferating cells suggests this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and would be useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases.

Alternatively, the protein product of this clone is useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions which include, but are not limited to Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia,

10

15

20

trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated

Expression of this gene product in regions of the brain suggests it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:79 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1495 of SEQ ID NO:79, b is an integer of 15 to 1509, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:79, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 70

The translation product of this gene shares sequence homology with insulinlike growth factor binding protein. Moreover, the protein has homology to the human
Slit-1 protein (See Genbank Accession No. gnllPIDld1036170 (AB017167)), which is
thought to play an integral role in neural development. In Drosophila embryogenesis,
the slit gene has been shown to play a critical role in CNS midline formation. Each
Slit gene encodes a putative secreted protein, which contains conserved proteinprotein interaction domains including leucine-rich repeats (LRR) and epidermal

growth factor (EGF)-like motifs, like that of the Drosophila protein. The Slit genes form an evolutionary conserved group in vertebrates and invertebrates, and the mammalian Slit proteins may participate in the formation and maintenance of the nervous and endocrine systems by protein-protein interactions.

5 In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: the EGF-like domain: CCCRLGLSGPKC (SEO ID NO:292); in addition to the following: RAFWGLGALQLLDLSANQLEAL (SEQ ID NO:293), HASGRRTGSADDGLQGRTGSGPPTAGAGGGGAAP (SEO ID NO:294), VSAAAGARLAPRAPGAPAGCRPMRGCAARAAARKSLVPVLPAGWRSGPAA 10 AARPGPRRLAHAPSAARSRAGPGAVARPLPRRHLAAAHGRGCGPAAARAGA GSGPGARRAARVPTAGRPPGTHVHTSGQSGAPRDPEGEALADTWAQTGQGD SSSNSSSSGRGRDQEGPRMGAAPPPPAPAVGGPLPVRPWSPSSAEPVLRPDAW I D (SEQ NO:295), TRPAAERAPRTTGSRDAQAAGLPPRVPGAGGLPPCGALPGR GLGRCCCCCCCLGLSGPKCRPGPRPRGPWAPRTAPRCARACREACQLSAL 15 SLPAVPPGLSLRLRALLLDHNRVRALPPGAFAGAGALORLDLRENGLHSVHV RAFWGLGALQLLDLSANQLEALAPGTFAPLRALRNLSLAGNRLARLEPAALG ALPLLRSLSLQDNELAALAPGLLGRLPALDALHLRGNPWGCGCALRPLCAWL RRHPLPASEAETVLCVWPGRLTLSPLTAFSDAAFSHCAQPLALRDLARGLHA 20 RAGLLPRQPGFLPGAGLWAHRLPCAPPPPPHRRPPPAETVQTRTPIPTPTAVPR PRTRGAPSAAAQA (SEQ I D NO:296), GCRPMRGCAARAARKSLVPVLPAGWRSGP AAAARPGPRRLAHAPSA (SEO ID NO:297), PGAVARPLPRRHLAAAHGRGCG PAAARAGA (SEQ ID NO:298), SGQSGAPRDPEGEALADTWAQTGQ (SEQ ID NO:299), 25 PPAPAVGGPLPVRPWSPSSAEPV (SEQ ID NO:300), APRTTGSRD AQAAGLPPRVPGAGGLP (SEQ ID NO:301), GPRPRGPWAPRTAPRCARACRE (SEQ ID NO:302), AVPPGLSLRLRALLLDHNRVRALPPGAFAGA (SEO ID NO:303), LGALQLLDLSANQLEALAPGTFAP (SEQ ID NO:304), PPGAFAGAG ALQRLDLRENGLHSVHVRAFWGLGALQ (SEQ ID NO:305), RNLSLAGNRLA 30 RLEPAALGALPLLRSLS (SEQ ID NO:306), LPALDALHLRGNPWGCGCALRP LCAW (SEQ ID NO:307), TVLCVWPGRLTLSPLTAFSDAAFSHCAQPLALRD

WO 99/47540 PCT/US99/05804

127

(SEQ ID NO:308), LHARAGLLPRQPGFLPGAGLWAHR (SEQ ID NO:309), and/or TVQTRTPIPTTAVPRPRTRGAPS (SEQ ID NO:310). Polynucleotides encoding these polypeptides are also encompassed by the invention.

It has been discovered that this gene is expressed primarily in a breast cancer cell line, MDA36.

5

10

15

25

30

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: neural, reproductive, and proliferative diseases and/or disorders, particularly breast cancer and degenerative conditions. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., neural, reproductive, and proliferative cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO.

177 as residues: Met-1 to Arg-10, Arg-64 to Ala-71, Gly-124 to Gly-131, Pro-189 to Arg-194, Val-223 to Gly-228.

The tissue distribution in a breast cancer cells and tissues and homology to insulin-like growth factor binding protien suggests that the protein product of this clone would be useful for diagnosis and treatment of breast cancer, and other forms of cancer. Moreover, the homology to the conserved human slit-1 protein suggests that the protein is useful in the treatment, diagnosis, and/or prevention of neural disorders, particularly developmental and degenerative conditions. Similarly, the protein is useful for the treatment and/or diagnosis of neurodegenerative disease states, behavioral disorders, or inflammatory conditions which include, but are not limited to Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia,

10

15

20

PCT/US99/05804

trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated

Expression of this gene product in regions of the brain suggests it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:80 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1095 of SEQ ID NO:80, b is an integer of 15 to 1109, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:80, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 71

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: HASGRPDRSSAPIGNSGLPCPDLEPLGGLQSKCRLCAPTE ARGLWSRSLCSDRCDTWRS (SEQ ID NO:311), and/or GLPCPDLEPLGGLQSK CRLCAPTEARGLW (SEQ ID NO:312). Polynucleotides encoding these polypeptides are also encompassed by the invention. This gene also maps to chromosome 1, and therefore can be used in linkage analysis as a marker for chromosome 1.

129

PCT/US99/05804

It has been discovered that this gene is expressed primarily in salivary gland and colon carcinoma.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: colon carcinoma and other digestive system or gastrointestinal diseases and/or disorders. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the digestive system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., digestive system, gastrointestinal, metabolic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, chyme, bile, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

10

15

20

25

30

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 178 as residues: Val-34 to Leu-39, Ser-64 to Cys-74, Ser-86 to Ser-95, Arg-128 to Ala-136.

The tissue distribution in salivary gland and colon carcinoma suggests that the protein product of this clone would be useful for the treatment and diagnosis colon cancer and other digestive system diseases and/or disorders, such as ulcers, and other proliferative conditions. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:81 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the

general formula of a-b, where a is any integer between 1 to 793 of SEQ ID NO:81, b is an integer of 15 to 807, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:81, and where b is greater than or equal to a + 14.

5

10

15

20

25

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 72

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: QEWESELGERRKPLQA (SEQ ID NO:313). Polynucleotides encoding these polypeptides are also encompassed by the invention.

It has been discovered that this gene is expressed primarily in 6 week old human embryos.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: embryological defects; aberrant development; aberrant cellular proliferation (e.g. cancers), and other developmentally related or proliferative diseases and/or disorders. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the developing human embryo, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., developmental, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, amniotic fluid, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

The tissue distribution in 6 week old human embryos suggests that the protein product of this clone would be useful for the diagnosis and/or treatment of defects in embryonic development. Elevated expression of this gene product in early 6 week human embryos suggests that this gene product plays a critical role in normal human development. Alternatively, this gene product may be involved in the pattern of cellular proliferation that accompanies early embryogenesis. Thus, aberrant

10

15

20

25

30

Expression of this gene product in tissues - particularly adult tissues - may correlate with patterns of abnormal cellular proliferation, such as found in various cancers. Moreover, this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and would be useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:82 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1029 of SEQ ID NO:82, b is an integer of 15 to 1043, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:82, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 73

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: CQSSNLIFFQFVNILFNLMMDILVDFSITKMPINSIFSLYF

CYEII (SEQ ID NO:314). Polynucleotides encoding these polypeptides are also encompassed by the invention.

It has been discovered that this gene is expressed primarily in 6 week old human embryo.

5

10

15

20

25

30

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: abnormal embryonic development; abnormal cellular proliferation; developmental defects, and other developmentally related or proliferative diseases and/or conditions. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the developing human embryo, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., developmental, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, amniotic fluid, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

The tissue distribution in 6 week old human embryo suggests that the protein product of this clone would be useful for the diagnosis and treatment of disorders of human embryonic development. Expression of this clone in developing embryos suggests that it plays a critical role in early human development. Alternatively, it may be involved in key cellular proliferation events that occur during embryogenesis. Therefore misexpression of this gene in adult tissues may lead to abnormal patterns of cellular proliferation and cancer. Moreover, expression within embryonic tissue and other cellular sources marked by proliferating cells suggests this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell

15

20

25

30

(3)

death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and would be useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:83 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1159 of SEQ ID NO:83, b is an integer of 15 to 1173, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:83, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 74

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: GPVWLFCFLTLCRKPSQLFSQENSCMDVAGGVTTCLPP WFSRGAPAQMSQWPPSSDHGAVRAGRDSRVGPVQPSHLTCEGGKEEREKNK KAEVNPPTGMGLANRIPRDDITLKLRNQGKLRTKENRTQSAKRHP (SEQ ID NO:315), VACKPENRTKTHFASSPACDGHALGGQVGFAICFLSCLFPPM (SEQ ID NO:316), and/or SHPMPNTPQKQLLFSEDNELLVSLRTGRKPTLQAALRVTG (SEQ ID NO:317). Polynucleotides encoding these polypeptides are also encompassed by the invention.

It has been discovered that this gene is expressed primarily in pleural cancer and endometrial tumors, and, to a lesser extent, in bone marrow & apoptotic T cells.

5

10

15

20

25

30

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: pleural cancer; endometrial tumors; hematopoietic disorders; immune dysfunction. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the lungs and immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., immune, hematopoietic, reproductive, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

The tissue distribution in pleural cancer and endometrial tumors indicates that the protein products of this clone are useful for the diagnosis and treatment of various reproductive cancers, including pleural cancer and endometrial tumors. In addition,

Expression of this gene product within T cells & bone marrow suggests that it may play a role in normal hematopoiesis. Therefore, this gene product may also be useful in the diagnosis and/or treatment of a variety of hematopoietic disorders, including defects in immune surveillance, inflammation, impaired immune function, and T cell lymphomas. Use of this gene product may be appropriate in situations designed to affect the proliferation, survival, and/or differentiation of various hematopoietic cell lineages, including blood stem cells.

Moreover, this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in

acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and would be useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:84 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1547 of SEQ ID NO:84, b is an integer of 15 to 1561, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:84, and where b is greater than or equal to a + 14.

20

5

10

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 75

The translation product of this gene shares low sequence homology with dreg-2, a gene product originally identified in Drosophila that shows an oscillating pattern of expression tied into a circadian clock rhythm.

In specific embodiments, polypeptides of the invention comprise the following a m i n o a c i d s e q u e n c e :

AHRLQIRLLTWDVKDTLLRLRHPLGEAYATKARAHGLEV

EPSALEQGFRQAYRAQSHSFPNYGLSHGLTSRQWWLDVVLQTFHLAGVQDA

QAVAPIAEQLYKDFSHPCTWQVLDGAEDTLRECRTRGLRLAVISNFDRRLEGI

LXGLGLREHFDFVLTSEAAGWPKPDPRIFQEALRLAHMEPVVAAHVGDNYL

CDYQGPRAVGMHSFLVVGPQALDPVVRDSVPKEHILPSLAHLLPALDCLEGS

T P G L (S E Q I D NO:319), EGDPRGRPRPRPLGPPPQLTLPTALXDILRQVRAPGLRLSRA
LEVGRKGSPIFKIQIYL (SEQ ID NO:318), IRLLTWDVKDTLLRLRHPLGEAYA
TKA (SEQ ID NO:320), LEQGFRQAYRAQSHSFPNYGLSHG (SEQ ID NO:321),
HLAGVQDAQAVAPIAEQLYKDFSHPC (SEQ ID NO:322), VLDGAEDTLRECR
TRGLRLAVIS (SEQ ID NO:323), REHFDFVLTSEAAGWPKPDPRIFQEA (SEQ ID NO:324), EPVVAAHVGDNYLCDYQGPRAVGMHSFL (SEQ ID NO:325),
and/or VVRDSVPKEHILPSLAHLLPALD (SEQ ID NO:326). Polynucleotides encoding these polypeptides are also encompassed by the invention.

5

10

15

20

25

30

It has been discovered that this gene is expressed primarily in tumors of the pancreas & thymus and to a lesser extent in a variety of fetal tissues, including fetal brain, liver, spleen, and kidney.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: pancreatic cancer; thymic cancer; disorders of fetal development; abnormal cellular proliferation; hematopoietic disorders. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the pancreas and immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., developmental, metabolic, immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, amniotic fluid, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

The tissue distribution in proliferative and developmental cells and tissues indicates that the protein products of this clone are useful for the diagnosis and treatment of cancers, particularly pancreatic and thymic cancer. Expression of this gene product within various fetal tissues also indicates that it is useful in the diagnosis and/or treatment of human developmental disorders. Taken together, the observation

that this gene product is expressed in cancers and in fetal tissues indicates that it plays a role in proliferation and/or differentiation events that are associated with early development. Misexpression of this gene product in adult tissues, therefore, may directly contribute to abnormal cellular proliferation and/or dedifferentiation that accompanies cancer. Finally,

Moreover, the expression of this gene product in fetal liver/spleen also suggests that it plays a role in hematopoiesis, and is useful in the diagnosis and/or treatment of a variety of disorders of the immune system. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

5

10

15

20

25

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:85 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1419 of SEQ ID NO:85, b is an integer of 15 to 1433, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:85, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 76

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: IRKLGPGLAPCSCRSGQVFPRV (SEQ ID NO:327).

Polynucleotides encoding these polypeptides are also encompassed by the invention.

It has been discovered that this gene is expressed primarily in frontal cortex, particularly derived from epileptic patients.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: epilepsy; neurodegenerative

WO 99/47540 PCT/US99/05804

138

diseases and disorders, particularly learning disorders. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain, CNS, and/or PNS, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., neural, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

10

15

20

25

30

ľ

The tissue distribution in frontal cortex tissue suggests that the protein product of this clone would be useful for the diagnosis and/or treatment of disorders of the brain and nervous system, particularly epilepsy. Moreover, the expression of this gene product suggests that it may play a role in various critical processes of the nervous system, including nerve survival, pathfinding, signal conductance, and/or synapse formation. It may have effects on various processes including homeostasis, learning, motor function, language, etc. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:86 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1363 of SEQ ID NO:86, b is an integer of 15 to 1377, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:86, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 77

In specific embodiments, polypeptides of the invention comprise the following a m i n o a c i d s e q u e n c e :

- KPLRMARPGGPEHNEYALVSAWHSSGSYLDSEGLRHQDD

 FDVSLLVCHCAAPFEEQGEAERHVLRLQFFVVLTSQRELFPRLTADMRRFRK

 PPRLPPEPEAPGSSAGSPGEASGLILAPGPAPLFPPLAAEVGMARARLAQLVRL

 AGGHCRRDTLWKRLFLLEPPGPDRLRLGGRLALAELEELLEAVHAKSIGDIDP

 QLDCFLSMTVSWYQSLIKVLLSRFPRAVAISKAQTWELSTWLR (SEQ. ID. NO.228)
- 10 NO:328), ARGTLELPTPLIAAHQLYNYVADHASSYHM (SEQ ID NO:329), SHCEWPGQG AQNTTSMPWCRHGTVLAPTWTLRDFDTR (SEQ ID NO:330), PLTTVSHLCPL
 - SLRVFTSHLDITAGHSHRDDTWVPIPALPLKHLRPPSSPFALGPWVSHPLMRW VQKLSHLHSNPGTGFSMGGKSAEKLKC (SEQ ID NO:331), STAARGAPGPGR AGGTPRSSPCQIHWGHRPPAGLLPIHDGLLVPEPDQSSPKPLPQSCRHFQSPDL GTQYLVALNQKFTDCSALVFWTPLRKDVSEVVFREALPVQPQDTRSPPAQLV STYHHLESVINTACFTLLDPPPLKGVDWTTECHCSLNHGPTRLPARGRTDQPF W A P G Q A R H (S E Q I D N O : 3 3 2) ,
- QRVRSTPRTPWALLGYSFRPPW (SEO ID NO:333), PGGPEHNEYALVSAWHSS GSYLDSEGLR (SEQ IDNO:334), DVSLLVCHCAAPFEEQGEAERHVLR (SEQ ID NO:335), RLTADMRRFRKPPRLPPEPEAPGSSAGS (SEQ ID NO:336), GEASGLI LAPGPAPLFPPLAAEVGM (SEO ID NO:337),

HQRLCNYVLRVCCPSLAAGTALPKHPQPLTHPGL

25 TLWKRLFLLEPPGPDRLRLGGRL (SEQ ID NO:338), and/or LAELEELLEAVHAKSIGDIDPQLDCFLS (SEQ ID NO:339). Polynucleotides encoding these polypeptides are also encompassed by the invention.

It has been discovered that this gene is expressed primarily in fetal liver/spleen and leukocytes, and to a lesser extent in a colon adenocarcinoma cell line.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for

diagnosis of the following diseases and conditions: hematopoietic disorders; immune dysfunction; colon cancer; colorectal adenocarcinoma. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system and colon, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., hematopoietic, immune, gastrointestinal, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 184 as residues: Leu-16 to Ser-23, Ser-38 to Pro-43, Gly-53 to Leu-60.

10

15

20

25

30

The tissue distribution in colon adenocarcinoma suggests that the protein product of this clone would be useful for the diagnosis and/or treatment of gastrointestinal diseases and/or disorders, particularly proliferative conditions. Expression of this gene product in fetal and proliferative cells and tissues suggests that it may be a marker cancers, and that it's misregulated expression may in fact contribute to the development or progression of the types of cancers dictated by its expression.

Similarly, the expression of this gene product in fetal liver/spleen - a primary site of early hematopoiesis - taken together with its expression in peripheral blood leukocytes suggests that this gene product may play a role in a variety of hematopoietic processes, including the survival, proliferation, activation, and/or differentiation of all blood cell lineages, including the totipotent hematopoietic stem cell. Such a gene product may therefore play a role in a variety of hematopoietic disorders including inflammation; immune dysfunction; defects in immune surveillance; and hematopoietic cancers and lymphomas. Similarly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent

WO 99/47540 PCT/US99/05804

of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA).

5

10

15

20

25

30

Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and would be useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:87 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1701 of SEQ ID NO:87, b is an integer of 15 to 1715, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:87, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 78

The gene encoding the disclosed cDNA is believed to reside on chromosome 20. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 20.

It has been discovered that this gene is expressed primarily in brain.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: neurodegenerative diseases and/or disorders. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s)

5

10

15

20

25

30

or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., neural, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder. This gene is believed to reside on chromosome 20, D20S111-D20S195. Polynucleotides corresponding to this gene are useful, therefore, as chromosome markers.

142

PCT/US99/05804

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 185 as residues: Met-1 to Tyr-6, Thr-38 to Ala-44.

The tissue distribution in brain tissue indicates that the protein products of this clone are useful for diagnosis and treatment of disorders of the central nervous system. Moreover, the protein product of this clone is useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions which include, but are not limited to Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception.

In addition, elevated expression of this gene product in regions of the brain suggests it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:88 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 403 of SEQ ID NO:88, b is an integer of 15 to 417, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:88, and where b is greater than or equal to a + 14.

10

15

20

25

30

1

1:

10

0

iı

FEATURES OF PROTEIN ENCODED BY GENE NO: 79

When tested against U937 cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activating sequence) promoter element. Thus, it is likely that this gene activates myeloid cells, and to a lesser extent, other immune and hematopoietic cells or cell types, through the JAK-STAT signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: FQLYFNPELIFKHFQIWRLITNFLFFGPVGFNFLFNMIFLY RYCRMLEEGSFRGRTADFVFMFLFGGFLMTLFGLFVSLVFLGQAFTIMLVYV WSRXNPYVRMNFFGLLNFQAPFLPWVLMGFSLLLGNSIIVDLLGIAVGHIYFF LEDVFPNQPGGIRILKTPSILKAIFDTPDEDPNYNPLPEERPGGFAWGEGQ SEQ I D NO : 3 4 0) , GVGQATVGKMAYQSLRLEYLQIPPVSRAYTTACVLTTAAVQLELITPF QLYFNPELIFKHFQIWRLITNFLFFGPVGFNFLFNMIFLYRYCRMLEEGSFRGR TADFVF (SEQ ID NO:341), LIFKHFQIWRLITNFLFFGPVGF (SEQ ID NO:342), FLYRYCRMLEEGSFRGRTADFVFMF (SEQ ID NO:343), LVFLGQAFTIMLVYV

WSRXNPYV (SEQ ID NO:344), VLMGFSLLLGNSIIVDLLGIA (SEQ ID NO:345), NQPGGIRILKTPSILKAIFDTPDED (SEQ ID NO:346), RLEYLQIPPVSRAYTTAC VLTTAAVQLE (SEQ ID NO:347), and/or RLITNFLFFGPVGFNFLFNMIFLYRYC RMLE (SEQ ID NO:348). Polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is believed to reside on chromosome 17. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 17.

5

10

15

20

25

30

It has been discovered that this gene is expressed primarily in smooth muscle, fetal brain, fetal liver and to a lesser extent in activated macrophage, colon cancer.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: developmental diseases, immune-related diseases, neural disorders, and vascular diseases and conditions. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system and central nervous system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., developmental, vascular, immune, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, amniotic fluid, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

The tissue distribution in fetal liver, macrophage, and fetal brain indicates that the protein products of this clone are useful for treating and diagosis of immune system-related diseases and CNS diseases. Moreover, the protein product of this clone is useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include bone marrow cell ex- vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as

WO 99/47540 PCT/US99/05804

infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Alternatively, the protein is useful in the detection, treatment, and/or prevention of vascular conditions, which include, but are not limited to, microvascular disease, vascular leak syndrome, aneurysm, stroke, atherosclerosis, arteriosclerosis, or embolism.

5

10

15

20

25

30

Moreover, the expression within fetal tissue and other cellular sources marked by proliferating cells, combined with the GAS biological activity, suggests this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and would be useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:89 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention

are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1153 of SEQ ID NO:89, b is an integer of 15 to 1167, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:89, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 80

The translation product of this gene shares sequence homology with proacrosin binding proteins (sp32) from non-human mammalian species. The binding of sp32 to proacrosin may be involved in packaging the acrosin zymogen into the 10 acrosomal matrix. See, for example, J Biol Chem. 1994 Apr 1; 269(13): 10133-10140, incorporated herein by reference. Accordingly, the inventors have termed the translation product of this gene human sp32 or "h-sp32". Contact of cells with supernatant expressing the product of this gene has been shown to increase the permeability of the plasma membrane of PMN to calcium. Thus it is likely that the 15 product of this gene is involved in a signal transduction pathway that is initiated when the product binds a receptor on the surface of the plasma membrane of both neutrophils, and to a lesser extent in other immune and hematopoietic cells. Thus, polynucleotides and polypeptides have uses which include, but are not limited to, 20 activating

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: **HASAGPDGSSPA** (SEO IDNO:349), ELLLEKPKPWQPPAAAPHRALLVLCYSIVENTCIITPTAKAWKYMEEEILGFG KSVCDSLGRRHMSTCALCDFCSLKLEQCHSEASLQRQQCDTSHKTPFAAPCL 25 PPRACPSATR (SEQ I D NO:350), LPGWGFPTKICDTDYIQYPNYCSFKSQQCLMR NRNRKVSRMRCLQNETYSALSPGKSEDVVLRWSQEFSTLTLGQFG (SEQ ID NO:351), SPVLLPAFPPLPVPLLALPVSAPLPACVLVSAPACAPLLAPACAL ALAPGFPGTRRIVGALPRCC (SEQ ID NO:352), LLVLCYSIVENTCIITPTAK AWKYMEEEILGFGKS (SEQ ID NO:353), and/or LKLEQCHSEASLQRQQC 30 DTSHKTPFA (SEQ ID NO:354). Polynucleotides encoding these polypeptides are

WO 99/47540 PCT/US99/05804

also encompassed by the invention. The gene encoding the disclosed cDNA is believed to reside on chromosome 12. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 12.

147

It has been discovered that this gene is expressed primarily in testis.

5

10

15

20

25

30

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: reproductive disorders. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive diseases, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., reproductive, testis, prostate, epidiymus, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, seminal fluid, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder. This gene is believed to map to chromosome 12 and is thought to be useful as a chromosome marker.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 187 as residues: Asp-27 to Ser-32, Pro-52 to Thr-58, Arg-63 to Asn-70, Gln-78 to Gly-83, Thr-107 to Asn-113, Thr-160 to Val-176, Ser-188 to Gly-241, Leu-248 to Pro-265, Tyr-302 to Gly-314.

The tissue distribution in testis, combined with the specific homology to the sp32 protein indicates that the protein products of this clone are useful for the diagnosis, treating, and/or prevention of reproductive diseases and/or disorders.

Moreover, polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of conditions concerning proper testicular function (e.g. endocrine function, sperm maturation), as well as cancer. Therefore, this gene product is useful in the treatment of male infertility and/or impotence. This gene product is also useful in assays designed to identify binding agents, as such agents (antagonists) are useful as male contraceptive agents.

15

20

25

30

Similarly, the protein is believed to be useful in the treatment and/or diagnosis of testicular cancer. The testes are also a site of active gene expression of transcripts that may be expressed, particularly at low levels, in other tissues of the body. Therefore, this gene product may be expressed in other specific tissues or organs where it may play related functional roles in other processes, such as hematopoiesis, inflammation, bone formation, and kidney function, to name a few possible target indications. The protein is useful in application and utility as a contraceptive, either directly or indirectly. Based upon the detected calcium flux activity, the protein may also be useful as an effect treatment for infertility (i.e. for inhibiting autoimmune disorders). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:90 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1878 of SEQ ID NO:90, b is an integer of 15 to 1892, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:90, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 81

The translation product of this contig has consistent sequence homology with a number of previously described viral tat proteins (see, for example, Stevens, et al., J. Virol. 64:3716-3725 (1990), which is hereby incorporated by reference, herein).

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: QVSGLILSLSCGMDGLALDGSPSPSPXTEKAGRCISQTSL (SEQ ID NO:355), QVSGLILSLSCGMDGLALDGSPSPSPXTEKAGRCISQTSLP GKWEV (SEQ ID NO:356), RASKTVPRMPPNWPAKMPCLCHIRTVEHLGTIS

10

15

20

25

30

SGAPGRPTGQQAARTYHICWIHPGQKIDSLPPSSQHPRSQQLAPGTWPSTSTT KPAEETLGSSASLPISQARKSEKCTFQPSPWXVRGKESHQVPAHPSHRTETES D HSPVRKPPSRGTRTGDFTVGDWSEAWLLELALL (SEQ ID NO:357), RMPPN WPAKMPCLCHIRTVEHLG (SEQ ID NO:358), GRPTGQQAARTYHICWIHPG QKIDS (SEQ ID NO:359), WPSTSTTKPAEETLGSSASLPISQA (SEQ ID NO:360), KSEKCTFQPSPWXVRGKESHQVP (SEQ ID NO:361), and/or KPPSRGTRTGDF TVGDWSEAWLLE (SEQ ID NO:362). Polynucleotides encoding these polypeptides are also encompassed by the invention.

It has been discovered that this gene is expressed almost exclusively in neutrophils.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of immune disorders. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder. In addition, molecules of the present invention can be used to regulate transcription and translation of genes in cells of the immune system, as well as in other cell types. Such transcriptional and translation regulation is useful for diagnosing and treating a number of disorders in which an alterred state of transcription and translation may be a factor in the disorder. Such disorders include many viral infections, particularly of immune cells, including HIV-1, HIV-2, human T-cell lymphotropic virus (HTLV)-I, and HTLV-II, as well as other DNA and RNA viruses such as herpes simplex virus (HSV)-1, HSV-2, HSV-6, cytomegalovirus (CMV), Epstein-Barr virus (EBV), herpes samirii, adenoviruses, rhinoviruses, influenza viruses, reoviruses, and the like. In addition, the ability to use molecules of the present invention to molecularly regulate the processes of transcription and

10

15

20

25

30

translation is useful in the diagnosis and treatment of many types of cancers, particularly those of the immune system, including ovarian cancer, breast cancer, colon cancer, cardiac tumors, pancreatic cancer, melanoma, retinoblastoma, glioblastoma, lung cancer, intestinal cancer, testicular cancer, stomach cancer, neuroblastoma, myxoma, myoma, lymphoma, endothelioma, osteoblastoma, osteoclastoma, osteoclastoma, osteosarcoma, chondrosarcoma, adenoma, and the like.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 188 as residues: Gln-2 to Trp-12, Ala-30 to Glu-35, Gln-42 to Ser-51.

The tissue distribution in neutrophils, combined with the homology to viral tat proteins suggests that the protein product of this clone is useful for the diagnosis and treatment of immune disorders, particularly viral infections and proliferative disorders. Further, since this clone has a high degree of sequence relatedness to factors which are involved in the regulation of transcription and translation, this clone is useful as a regulator of such processes. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:91 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 509 of SEQ ID NO:91, b is an integer of 15 to 523, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:91, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 82

The translation product of this contig has clear sequence identity with a number of thioredoxins and endoplasmic reticulum resident proteins (see, for

10

15

20

25

example, Shorrosh and Dixon, Plant J. 2:51-58 (1992), which is hereby incorporated by reference, herein).

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: PCADCLSAWA (SEQ ID NO:363). Polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is believed to reside on chromosome 5. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 5.

It has been discovered that this gene is expressed primarily in adipocytes and striatum depression, and in lower abundance in prostate, whole brain, fetal liver, and spleen.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: Prostate cancer, CNS diseases, immune disorders. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., neural, hematopoietic, immune, and cancerous and wounded tissues) or bodily fluids (e.g., seminal fluid, amniotic fluid, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder. Since the translation product of this clone has a high degree of sequence relatedness to many thioredoxins, it can be used as a food additive to improve flour quality or to suppress the anti-nutritional effects of leguminous plants. Molecules of the present invention can further used to inactivate toxins, for example, bee or snake venom.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 189 as residues: Trp-43 to Ala-49, Pro-68 to Ala-74, Glu-100 to Gly-111, Glu-120 to Asn-125, Pro-141 to Ala-154, Asp-157 to Lys-171, Cys-177 to Ile-182, Ser-248 to Leu-253, Thr-280 to Glu-285, Gly-353 to Val-359.

WO 99/47540 PCT/US99/05804

152

The tissue distribution in whole brain suggests that the protein product of this clone would be useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions which include, but are not limited to Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated

5

10

15

20

25

30

Expression of this gene product in regions of the brain suggests it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. The secreted protein can also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, and as nutritional supplements. It may also have a very wide range of biological activities. Typical of these are cytokine, cell proliferation/differentiation modulating activity or induction of other cytokines; immunostimulating/immunosuppressant activities (e.g. for treating human immunodeficiency virus infection, cancer, autoimmune diseases and allergy); regulation of hematopoiesis (e.g. for treating anemia or as adjunct to chemotherapy); stimulation or growth of bone, cartilage, tendons, ligaments and/or nerves (e.g. for treating wounds, stimulation of follicle stimulating hormone (for control of fertility); chemotactic and chemokinetic activities (e.g. for treating infections, tumors); hemostatic or thrombolytic activity (e.g. for treating hemophilia, cardiac infarction etc.); anti-inflammatory activity (e.g. for treating septic shock, Crohn's disease); as antimicrobials; for treating psoriasis or other hyperproliferative diseases; for regulation of metabolism, and behavior. Also contemplated is the use of the corresponding nucleic acid in gene therapy procedures. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

10

15

20

25

30

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:92 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1368 of SEQ ID NO:92, b is an integer of 15 to 1382, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:92, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 83

When tested against TF-1 cell lines, supernatants removed from cells containing this gene activated the ISRE (interferon-sensitive responsive element) promoter element. Thus, it is likely that this gene activates myeloid cells, and to a lesser extent, in immune and hematopoietic cells or tissues, through the JAK-STAT signal transduction pathway. ISRE is a promoter element found upstream in many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the ISRE element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: HASGYLCIVLL (SEQ ID NO:364). Polynucleotides encoding these polypeptides are also encompassed by the invention.

It has been discovered that this gene is expressed exclusively in Rejected Kidney.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: kidney and other urinary tract

WO 99/47540 PCT/US99/05804

154

disorders and disorders related to, or resulting from, transplantation. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and renal systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., renal, kidney, urogenital, immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder. Molecules of the present invention are particularly useful in the diagnosis and treatment of disorders related to transplantation, particularly kidney transplantation.

5

10

15

20

25

30

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 190 as residues: Asn-49 to Gln-54, Glu-150 to Asp-159.

The tissue distribution in rejected kidney tissue suggests that the protein product of this clone would be useful for diagnosis and treatment of disorders related to or resulting from rejection of transplanted organs, particularly the kidney.

Moreover, the protein product of this clone could be used in the treatment and/or detection of kidney diseases including renal failure, nephritus, renal tubular acidosis, proteinuria, pyuria, edema, pyelonephritis, hydronephritis, nephrotic syndrome, crush syndrome, glomerulonephritis, hematuria, renal colic and kidney stones, in addition to Wilm's Tumor Disease, and congenital kidney abnormalities such as horseshoe kidney, polycystic kidney, and Falconi's syndrome. Considering the tissue distribution and detected ISRE biological activity, the protein is useful in modulating the immune response to aberrant kidney proteins, including autoantigens and aberrant proteins which are often present in degenerative and proliferative conditions. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:93 and may have been publicly available prior to conception of

15

20

25

the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1733 of SEQ ID NO:93, b is an integer of 15 to 1747, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:93, and where b is greater than or equal to a + 14.

10 FEATURES OF PROTEIN ENCODED BY GENE NO: 84

The translation product of this gene shares sequence homology with the conserved MAL and plasmolipin protein (Magyar, et al, Gene 189:269-275 (1997); See Genbank Accession No.gnllPIDle183885), which are thought to be important in modulating T cell function, and proper CNS function, respectively. When tested against Jurkat cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activating sequence) promoter element. Thus, it is likely that this gene activates myeloid cells, and to a lesser extent, immune or hematopoietic cells and tissues, through the JAK-STAT signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: NSARAARAEIVLGLLVWTLIAGTEYFRVPAFGWV (SEQ ID NO:365). Polynucleotides encoding these polypeptides are also encompassed by the invention.

It has been discovered that this gene is expressed primarily in T cells.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of immune, hematopoietic, and neural diseases and/or disorders. Similarly,

WO 99/47540

5

10

15

20

25

30

polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder. Nucleic acids of the present invention are useful as probes for detecting traumatic and pathological changes in the central and peripheral nervous systems. Molecules of the present invention may be involved in regulating the growth of Schwann cells and other neural cells. Molecules of the present invention are also useful as modulators of the interaction between Schwann cells and other neural cells and the extracellular matrix and is therefore useful for the therapeutic intervention in nerve damage primarily by facilitating regeneration of damaged axons and regenerating nerve cells in damaged nervous system tissues.

156

PCT/US99/05804

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 191 as residues: Ser-58 to His-64.

The tissue distribution in T-cells, combined with the homology to the MAL and plasmolipin proteins and the detected GAS biological activity suggests that the protein product of this clone would be useful for the diagnosis and treatment of immune disorders including, but not limited to, AIDS and other immunodeficiencies. Morever, the expression of this gene product suggests a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne,

10

15

20

25

30

neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosis, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types.

The secreted protein can also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, and as nutritional supplements. It may also have a very wide range of biological activities. Typical of these are cytokine, cell proliferation/differentiation modulating activity or induction of other cytokines; immunostimulating/immunosuppressant activities (e.g. for treating human immunodeficiency virus infection, cancer, autoimmune diseases and allergy); regulation of hematopoiesis (e.g. for treating anemia or as adjunct to chemotherapy); stimulation or growth of bone, cartilage, tendons, ligaments and/or nerves (e.g. for treating wounds, stimulation of follicle stimulating hormone (for control of fertility); chemotactic and chemokinetic activities (e.g. for treating infections, tumors); hemostatic or thrombolytic activity (e.g. for treating hemophilia, cardiac infarction etc.); anti-inflammatory activity (e.g. for treating septic shock, Crohn's disease); as antimicrobials; for treating psoriasis or other hyperproliferative diseases; for regulation of metabolism, and behavior. Also contemplated is the use of the corresponding nucleic acid in gene therapy procedures. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are

WO 99/47540

related to SEQ ID NO:94 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 586 of SEQ ID NO:94, b is an integer of 15 to 600, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:94, and where b is greater than or equal to a + 14.

158

PCT/US99/05804

10

15

20

25

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 85

The translation product of this clone has sequence identity to a protein tyrosine kinase reported by Oates and Wilks (The Worm Breeders Gazette 14:87-87 (1995), which is hereby incorporated by reference herein). The gene encoding the disclosed cDNA is believed to reside on chromosome 2. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 2.

It has been discovered that this gene is expressed primarily in cerebellum, adult brain, retina, spinal cord, and kidney cortex.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: neural, visual, and renal diseases and/or disorders. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the CNS, retina, and kidney cortex. Expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., neural, visual, renal, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

WO 99/47540

5

10

15

20

The tissue distribution in cerebellum, adult brain, and spinal cord tissue suggests that the protein product of this clone would be useful for the diagnosis and treatment of neural diseases and disorders. The protein product of this clone is useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions which include, but are not limited to Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated

159

PCT/US99/05804

Expression of this gene product in regions of the brain suggests it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Moreover, the protein product of this clone could be used in the treatment and/or detection of kidney diseases including renal failure, nephritus, renal tubular acidosis, proteinuria, pyuria, edema, pyelonephritis, hydronephritis, nephrotic syndrome, crush syndrome, glomerulonephritis, hematuria, renal colic and kidney stones, in addition to Wilm's Tumor Disease, and congenital kidney abnormalities such as horseshoe kidney, polycystic kidney, and Falconi's syndrome. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly
available and accessible through sequence databases. Some of these sequences are
related to SEQ ID NO:95 and may have been publicly available prior to conception of
the present invention. Preferably, such related polynucleotides are specifically
excluded from the scope of the present invention. To list every related sequence
would be cumbersome. Accordingly, preferably excluded from the present invention
are one or more polynucleotides comprising a nucleotide sequence described by the
general formula of a-b, where a is any integer between 1 to 572 of SEQ ID NO:95, b

is an integer of 15 to 586, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:95, and where b is greater than or equal to a + 14.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 86

10

15

20

25

30

The translation product of this clone has homology to trkB, and it is thought that the protein of the present invention is a novel novel neural receptor protein-tyrosine kinase, a trkB homolog (See for example,). This protein is likely to be derived from a gene for a ligand-regulated receptor closely related to the human trk oncogene. Northern (RNA) analysis showed that the trkB gene is expressed predominantly in the brain and that trkB expresses multiple mRNAs, ranging from 0.7 to 9 kb. Hybridization of cerebral mRNAs with a variety of probes indicates that there are mRNAs encoding truncated trkB receptors.

In specific embodiments, polypeptides of the invention comprise the sequence PCSPPDSPPLPGAFVWRVLWVC (SEQ ID NO:366). Polynucleotides encoding this polypeptide are also encompassed by the invention.

It has been discovered that this gene is expressed primarily in breast cancer, colon tumor, and B-cell lymphoma.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: breast cancer, colon tumor, B-cell lymphoma. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., neural, gastrointestinal, immune, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

WO 99/47540

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 193 as residues: Ser-29 to Asn-40.

The tissue distribution in proliferative cells and tissues suggests that the protein product of this clone would be useful for the treatment, detection, and/or prevention of cancer, particularly in the indicated tissues. The expression within 5 cellular sources marked by proliferating cells suggests this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. 10 Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, 15 detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and would be useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases.

Alternatively, the homology to the trkB protein suggests the protein product of 20 this clone is useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions which include, but are not limited to Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal 25 cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain suggests it plays a role in 30 normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal

differentiation or survival. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:96 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 788 of SEQ ID NO:96, b is an integer of 15 to 802, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:96, and where b is greater than or equal to a + 14.

15

20

25

5

10

FEATURES OF PROTEIN ENCODED BY GENE NO: 87

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: ARACFAYNGVCSEGRCWDSHFHGSV (SEQ ID NO:367), MSNMGKIPSLSLHIPINKYICSRIPKFIQKVNKSTVLQICLKRQIILNKNKMSDH SKIGKANLVQIDIHSLGIVETGCVPSKRYCTLLTEQSGFPFLSHP (SEQ ID NO:368),

MAGCCLKLFGVLSLCFLCGLISIERVICNPVSADFQVSTFCQRHCLLR SKVMFXIKGXTATIEVINENCTLVAAPPIGFPIXFL (SEQ ID NO:369), MSDHS KIGKANLVQIDIHSLGIVETGCVPSKRYCTLLTEQSGFPFLSHP (SEQ ID NO:370), MAGCCLKLFGVLSLCFLCGLISIERVICNPVSADFQVSTFCQRHCL LRSK (SEQ ID NO:371), VMFXIKGXTATIEVINENCTLVAAPPIGFPIXFL (SEQ ID NO:372). Polynucleotides encoding these polypeptides are also encompassed by the invention.

It has been discovered that this gene is expressed primarily in dendritic cells, and smooth muscle.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: immune, hematopoietic, and vascular diseases and/or disorders. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, expression of this gene at significantly higher or lower levels may be detected in certain tissues (e.g., immune, hematopoietic, smooth muscle vascular, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

5

10

15

20

25

30

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 194 as residues: Asp-40 to Ser-52.

The tissue distribution in dendritic cells suggests that the protein product of this clone would be useful for immune disorders.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:97 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1212 of SEQ ID NO:97, b is an integer of 15 to 1226, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:97, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 88

WO 99/47540 PCT/US99/05804

164

The translation product of this gene shares sequence homology with androgendependant expressed protein from golden hamster hair follicles which is thought to be important in regulating the secretions from glands in the skin (See GenBank Accession No. gil191315).

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: PTEGRQKVLKTFTVPRSALAMTKTSTCIYHFLVLSWYTF LNYYISQEGKDEVKPKILANGARWKY (SEQ ID NO:373), PTEGRQKVLKTF TVPRSALAMTKT (SEQ ID NO:375), PRSALAMTKTSTCIYHFLVLSWYTFLN YYISQEGK (SEQ ID NO:374), and/or FLNYYISQEGKDEVKPKILANGARWKY (SEQ ID NO:376). Polynucleotides encoding these polypeptides are also encompassed by the invention.

5

10

15

20

25

30

It has been discovered that this gene is expressed primarily in lung, colon cancer, and testis.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: disorders of secretory cells including cells in the lung, colon, testis and the skin. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the secretory epithelial cells in the lung, intestine, testis and skin, expression of this gene at significantly higher or lower levels may be detected in certain tissues (e.g., cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 195 as residues: Val-21 to Asp-30, Pro-101 to Thr-109.

The tissue distribution and homology to androgen regulated protein suggests that the protein product of this clone would be useful for treating disorders that involve highly secretory cells including those in the colon, testis, and skin. It may be useful for diagnosing disorders such as colon, lung, or testicular cancer and may be

used to treat pulmonary conditions in patients with compromised respiratory function. In addition, the polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of conditions concerning proper testicular function (e.g. endocrine function, sperm maturation), as well as cancer. Therefore, this gene product is useful in the treatment of male infertility and/or impotence. This gene product is also useful in assays designed to identify binding agents, as such agents (antagonists) are useful as male contraceptive agents.

5

10

15

20

25

Similarly, the protein is believed to be useful in the treatment and/or diagnosis of testicular cancer. The testes are also a site of active gene expression of transcripts that may be expressed, particularly at low levels, in other tissues of the body. Therefore, this gene product may be expressed in other specific tissues or organs where it may play related functional roles in other processes, such as hematopoiesis, inflammation, bone formation, and kidney function, to name a few possible target indications. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:98 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1106 of SEQ ID NO:98, b is an integer of 15 to 1120, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:98, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 89

The translation product of this gene shares sequence homology with dec-205 a transmembrane protein which is thought to be important in antigen presentation in dendritic cells and T-cells.

WO 99/47540 PCT/US99/05804

166

It has been discovered that this gene is expressed primarily in macrophage, dendritic cells, lung and ulcerative colitis.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: inflammatory diseases such as ulcerative colitis. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues (e.g., cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

10

15

20

25

30

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 196 as residues: Asp-30 to Arg-36, Gln-59 to Val-65.

The distribution in macrophage, dendritic cells, lung and ulcerative colitis tissues, and homology to antigen presenting receptors suggests that the protein product of this clone would be useful for modulating the immune response in both acute and chronic inflammatory conditions. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:99 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2582 of SEQ ID NO:99, b is an integer of 15 to 2596, where both a and b correspond to the positions of

PCT/US99/05804

nucleotide residues shown in SEQ ID NO:99, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 90

5

10

15

20

25

30

This gene maps to chromosome 22 and therefore polynucleotides of the present invention can be used in linkage analysis as a marker for chromosome 22.

In specific embodiments, polypeptides of the invention comprise the sequence FKDQLVYPLLAFT (SEQ ID NO:377) and/or RQALNLPDVFGLV (SEQ ID NO:379). Polnucleotides encoding these polypeptides are also encompassed by the invention.

It has been discovered that this gene is expressed primarily in fetal spleen and liver as well as cd34 positive cells and to a lesser extent in several tissues suggesting a presence in blood or blood forming tissues.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: developmental defects in the blood and blood forming cells. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., fetal spleen and liver as well as cd34 positive cells, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 197 as residues: Gln-54 to Gly-61, Asn-79 to Leu-91, Glu-99 to Thr-105, Pro-120 to Gln-126, Pro-128 to Phe-134, Arg-150 to Arg-156, Arg-160 to Arg-170.

The tissue distribution in fetal spleen and liver as well as cd34 positive cells suggests that the protein product of this clone would be useful for treating disorders in

the development, proliferation, or regulation of blood forming cells including diseases such as lymphomas, granulomas, leukemias, and in the preservation and or replenishment of stem cells in the blood.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:100 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1006 of SEQ ID NO:100, b is an integer of 15 to 1020, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:100, and where b is greater than or equal to a + 14.

15

20

25

30

10

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 91

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: ATASHDLLLF (SEQ ID NO:379), MSINICLMQSKTQGSCQ YLLLPHPVPIILKVSTVFSLLSLFRLLFLSFCPHPKKCSYLLKYYGPLEGHKTLX YLRTNLGVIQPPLRMYAAEDCNGIG (SEQ ID NO:380), MSINICLMQSKTQG SCQYLLLPHPVPIILKVSTVFSLLSLFRLLFL (SEQ ID NO:381), and/or SFCPHPK KCSYLLKYYGPLEGHKTLXYLRTNLGVIQPPLRMYAAEDCNGIG (SEQ ID NO:382). Polynucleotides encoding these polypeptides are also encompassed by the invention.

It has been discovered that this gene is expressed primarily in T cells, fetal heart and chronic lymphocytic leukemia and to a lesser extent in kidney, lung, and 16 week embryos.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: disorders of the blood including abnormalities in T cell function or blood cell proliferation such as leukemia.

WO 99/47540

10

15

20

25

30

Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., T cells, fetal heart and chronic lymphocytic leukemia, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

169

PCT/US99/05804

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 198 as residues: Leu-45 to Val-50.

The tissue distribution in T cells, fetal heart and chronic lymphocytic leukemia suggests that the protein product of this clone would be useful for treating abnormalities of the blood particularly those involving T-cells and the abnormal proliferation of blood cells such as lymphocytic leukemia. In addition, it suggests the protein product of this clone is useful for the diagnosis and treatment of a variety of immune system disorders. Morever, the expression of this gene product suggests a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury. demyelination, systemic lupus erythematosis, drug induced hemolytic anemia,

WO 99/47540 PCT/US99/05804

170

rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types.

5

10

15

20

25

30

The expression in fetal heart tissue would suggest a useful role for the protein product in developmental abnormalities, fetal deficiencies, pre-natal disorders and variouswould-healing models and/or tissue trauma. The tissue distribution in kidney suggests the protein product of this clone could be used in the treatment and/or detection of kidney diseases including renal failure, nephritus, renal tubular acidosis, proteinuria, pyuria, edema, pyelonephritis, hydronephritis, nephrotic syndrome, crush syndrome, glomerulonephritis, hematuria, renal colic and kidney stones, in addition to Wilm's Tumor Disease, and congenital kidney abnormalities such as horseshoe kidney, polycystic kidney, and Falconi's syndrome.

In addition, the tissue distribution in embryonic tissue suggests the protein product of this clone is useful for the diagnosis, detection, and/or treatment of developmental disorders. Expression within embryonic tissue and other cellular sources marked by proliferating cells suggests this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and would be useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. Protein, as well as, antibodies

20

25

30

directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:101 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1506 of SEQ ID NO:101, b is an integer of 15 to 1520, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:101, and where b is greater than or equal to a + 14.

15 FEATURES OF PROTEIN ENCODED BY GENE NO: 92

The translation product of this gene shares sequence homology with ctg4 which is a glutamine repeat containing gene thought to be a candidate genetic disease locus.

In specific embodiments, polypeptides of the invention comprise the sequence KEEDDDTERLPSKCEVCKLLSTE (SEQ ID NO:383 and 384) LQAELSRTGRSR EVLELGQ (SEQ ID NO:385 and 386), RQAVIVCRRRFV (SEQ ID NO:387), PPRWAHPKAPEGSPDPPSPPSALGLSVLPWSDSDPWHISVSPCAQREHYSPGS AHINSLRPLPALSLKRCKARVSSSCLYPAPAPAPAPLEIDRCDSVPPVALCSAA YTLRICWASVLCHRPPPSTSQPKPRARPKKGKAIFPTAQVP (SEQ ID NO:388), PPRWAHPKAPEGSPDPPSPPSALGLSVLPWSDSDPWHISVSPCAQREHYSPGS AHINSLRPLPALSLKRCK (SEQ ID NO:389), and/or ARVSSSCLYPAPAPAPAPL EIDRCDSVPPVALCSAAYTLRICWASVLCHRPPPSTSQPKPRARPKKGKAIFPT AQVP (SEQ ID NO:390). Polynucleotides encoding these polypeptides are also encompassed by the invention.

It has been discovered that this gene is expressed in several tissues including lung, heart, kidney, adrenal gland, smooth muscle, cerebellum, and embryonic tissue.

10

15

20

25

30

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: inherited developmental disorders possibly with a neuropsychiatric component. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 199 as residues: Lys-25 to Ser-36, Ser-53 to Glu-60, Thr-70 to Arg-75, Arg-111 to Thr-119, Glu-161 to Leu-189.

The tissue distribution and homology to glutamine repeat family member CTG4 suggests that the protein product of this clone would be useful for identifying and treating specific diseases related to nucleotide triplet expansion. The tissue distribution in embryonic tissue suggests the protein product of this clone is useful for the diagnosis, detection, and/or treatment of developmental disorders. The relatively specific expression of this gene product during embryogenesis suggests it may be a key player in the proliferation, maintenance, and/or differentiation of various cell types during development. It may also act as a morphogen to control cell and tissue type specification. Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. Expression within embryonic tissue and other cellular sources marked by proliferating cells suggests this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are

WO 99/47540

related to SEQ ID NO:102 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1292 of SEQ ID NO:102, b is an integer of 15 to 1306, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:102, and where b is greater than or equal to a + 14.

10

15

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 93

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: EEKLFTSAPGRDFWVMGETRDGNEEN (SEQ ID NO:391). Polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is believed to reside on chromosome 16. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 16.

It has been discovered that this gene is expressed primarily in cancerous and fetal tissue.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: cancer, developmental anomalies or fetal deficiencies. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system and developing fetus, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., developmental, reproductive, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene

WO 99/47540

5

10

15

20

25

30

PCT/US99/05804

174

expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 200 as residues: Met-1 to Ser-6.

The tissue distribution in fetal tissue suggests that the protein product of this clone would be useful for the treatment and diagnosis of developmental anomalies or fetal deficiencies. In addition to fetal tissue, expression in a variety of cancerous tissues suggests a role in the treatment and diagnosis of uncontrolled cell proliferation and/or differentiation (e.g. cancer). Moreover, the expression within embryonic tissue and other cellular sources marked by proliferating cells suggests this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders.

Similarly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and would be useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:103 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the

PCT/US99/05804

general formula of a-b, where a is any integer between 1 to 771 of SEQ ID NO:103, b is an integer of 15 to 785, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:103, and where b is greater than or equal to a + 14.

5

10

15

20

25

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 94

The gene encoding the disclosed cDNA is believed to reside on chromosome 10. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 10.

This gene is expressed primarily in hypothalamus, T-cells, and adipose tissue.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: immune (e.g. immunodeficiencies, autoimmunities, inflammation, leukemias & lymphomas) and neurological (e.g. Alzheimer's disease, dementia, schizophrenia) disorders. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders

immune systems, expression of this gene at significantly higher or lower levels may be detected in certain tissues (e.g., immune, neural, metabolic, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder. The tissue distribution suggests that the protein product of this clone would be useful in the intervention or detection of pathologies associated with the hematopoietic and immune systems, such as anemias (leukemias). In addition, the

of the above tissues or cells, particularly of the central nervous, hematopoietic and

expression in brain (including fetal) might suggest a role in developmental brain defects, neuro-degenerative diseases or behavioral abnomalities (e.g. schizophrenia,

Alzheimer's, dementia, depression, etc.).

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 201 as residues: Phe-64 to Gly-77, Pro-83 to Asp-99.

PCT/US99/05804

The tissue distribution in hypothallamus suggests the protein product of this clone is useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions which include, but are not limited to Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, 5 Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in 10 feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain suggests it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. This gene product may be involved in the regulation of cytokine production, 15 antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosis, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues.

20

25

30

Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Moreover, the protein

product of this clone is useful for the diagnosis, prevention, and/or treatment of various metabolic disorders which include, but are not limited to, Tay-Sachs disease, phenylkenonuria, galactosemia, hyperlipidemias, porphyrias, and Hurler's syndrome. The protein is useful in the treatment and/or prevention of neurodegenerative conditions, particularly those which occur secondary to aberrant fatty acid metabolism (i.e. defects which affect the synthesis and integrity of the myelin sheath). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:104 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2001 of SEQ ID NO:104, b is an integer of 15 to 2015, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:104, and where b is greater than or equal to a + 14.

20

25

30

5

10

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 95

The translation product of this gene was shown to have homology to the murine leucine-rich repeat protein (See Genbank Accession No. gil2880079), which is thought to be important in neural development.

In specific embodiments, the polypeptides of the invention comprise the sequence:QKPTFALGELYPPLINLWEAGKEKSTSLKVKATVIGLPTNMS (SEQ ID NO:392). Polynucleotides encoding this polypeptide are also encompassed by the invention. The gene encoding the disclosed cDNA is believed to reside on chromosome 7. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 7.

10

15

20

25

30

It has been discovered that this gene is expressed primarily in T-cells and brain.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: immunodeficiency, tumor necrosis, infection, lymphomas, auto-immunities, cancer, inflammation, anemias (leukemia) and other hematopoeitic disorders, neurological diseases of the brain such as depression, schizophrenia, Alzheimer's disease, Parkinson's disease, Huntington's disease, dementia and specific brain tumors. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain and immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues or cell types (e.g., neural, immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid, serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 202 as residues: Met-24 to Gly-29, Ala-57 to Thr-63.

The tissue distribution in T-cells suggests that the protein product of this clone would be useful for the diagnosis and treatment of immune disorders including: leukemias, lymphomas, auto-immunities, immunodeficiencies (e.g. AIDS), immunosupressive conditions (transplantation) and hematopoeitic disorders. In addition this gene product may be applicable in conditions of general microbial infection, inflammation or cancer. The expression in brain, combined with the homology to the leucine-rich repeat protein suggests that the protein product of this clone would be useful for the treatment and diagnosis of developmental, degenerative and behavioral conditions of the brain and nervous system, such as depression, schizophrenia, Alzheimer's disease, Parkinson's disease, Huntington's disease, Tourette Syndrome, mania, dementia, paranoia, addictive behavior, obsessive-compulsisve disorder and

sleep disorders. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:105 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 353 of SEQ ID NO:105, b is an integer of 15 to 367, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:105, and where b is greater than or equal to a + 14.

10

		#			Ŀ	Т	_	T		T		1		T		T_		Т		T. ~		Т		Т	
L		Last	AA		ORF	74		2	}	371	}	57		113		151		55)	236	, 	4	?	64	,
		First	AA of	Secreted		35		28	}	33)	33	!	21	ı 	21	ł	37		21	i 	43	!	31	i !
	First Last	AA	Jo	Sig	Pep	34		77	i	32	1	32		20		20		36		20		42	!	38	
		AA	Jo	Sig	Pep	-	ı	-	•	-		-		-		-				-		-		-	
	AA	SEQ	Ω	ÖN	>	108		109	·	110		203		111		112		113		114		115		116	
5' NT	Jo	_	AA of		Pep	94		58)	397		193		368		164		43		106		268	•	175	
		5' NT	Jo	Start	Codon	94		58		397		193		368		164		43		106		268		175	
	5' NT 3' NT	Jo	Clone	Seq.	'	2343		1083		2107		1889		1256		759		1810		296		1130	 -	883	
	5'NT	Jo	Total Clone Clone	Sed.	1	_		_		197		_		208		_		-		-		-		-	
			Total	NT	Seq.	2343		1177		2107		1889		1262		759		1810		1052		1130		883	
	NT	SEQ	Ω	NO:	×	=		12		13		106		14		15		16		17		18		19	
					Vector	pSport1	•	Uni-ZAP XR		pCMVSport	3.0	pCMVSport	3.0	pCMVSport	3.0	pCMVSport	3.0	Lambda ZAP	II	Uni-ZAP XR		Uni-ZAP XR		Uni-ZAP XR	
		ATCC	Deposit	Nr and	Date	209628	02/12/98	879607	02/12/98	879607	02/17/98	879607	02/12/98	209628	02/12/98	209628	02/12/98	209628	02/12/98	209628	02/12/98	209628	02/12/98	209628	02/12/98
					Clone ID	HKGCR51		HPMDK28		HLDCD04		HLDCD04		HLDON23		HLDRM43		HLQAM28		HLTDE74		HLTFA64		HMCFY13	
				Gene	No.	-		7		3		ε.		4		2		9		7		∞		6	

Last	AA	jo	2	117	49	42	47	139	91	346	233	53	81	41
First	AA of	Secreted	rormon	21	15	21	25	44	24	35	35	24	35	22
Last AA	Jo	Sig	Leb	70	14	20	24	43	23	34	34	23	34	21
First Last AA AA	Jo	Sig	rep	_	_	1	1	-	_	-	-	-	-	1
AA SEQ	Ð	: 	I	117	118	119	120	121	122	123	124	125	126	127
5° NT of First	AA of	<u> </u>		237	185	57	84	34	901	49	228	35	57	163
5° NT	Jo	Start	Conoil	237	185	57	84	34	106	49	228	35	57	163
	Clone	Seq.		686	495	2317	1726	529	1755	1451	1212	1112	748	778
5' NT 3' NT of	Total Clone Clone	Seq.		169	-	-	-		-	-	28	-	-	-
	Total	L Z	350	686	495	2317	1726	529	1755	1751	1212	1112	748	778
NT SEQ	Ω	;; ;;	4	20	21	22	23	24	25	26	27	28	29	30
	-	Voston	VCCIOI	pSport1	Uni-ZAP XR	Uni-ZAP XR	Uni-ZAP XR	pSport1	Uni-Zap XR	pSport1	pBluescript	Uni-ZAP XR	Uni-ZAP XR	Uni-ZAP XR
ATCC	Deposit	Nr and	Dalle	209628 02/12/98	209628	209628	209628 02/12/98	209628 02/12/98						
	,	cDNA Clere ID	Ciolic ID	HMMBD35	НМОСУ03	HMSBX84	HMSKI86	HMVBS81	нммев02	HMZAD77	HNFIY77	HNHEK85	нинеп93	НОДАН74
	(Gene		10	1	12	13	14	15	16	17	18	19	20

Γ		Last	AA	of L	<u></u>	 6	Ę	,	2	077	1	/11		305 	 ;	 -	;		Ş	4 4	0	 2		0/7		00	
		First L			ᆰ	76		4	†	77	\dagger	×2	\dagger	32	;	3		3.	!	_	6	87		32	;	/7	
一	Last	AA			_ †	25	1	43	1	17	!	17	1	31	1	74	1	30	1	91	į	/7		31	į	97	
r	First 1	AA	of		핅		1.	_	1	-	1			_	1		1	_]	-		_					
r	AA	SEQ	<u> </u>		-	128	19	129	1	130		131		132		133		134		135	;	136		137		138	
S' NT		First		Signal NO:	-	229		225		104		411		11		37		138		215		318		146		 48	
	<u> </u>	S' NT	Jo	Start	Codon	229		225		104		411		77		37		138		215		318		146		84	
	L'A	_		Seq.	-	1324	1	739		1462		2815		1076		1217		1282		559		803		1510		1095	
r	5' NT 3' NT	Jo	Clone Clone	Seq.	—			6	_	73				_				_		_		215		-			
	<u>v</u>	-	Total (K	Seq.	1324		739		1462		2815		1078		1217		1282		559		803		1510		1095	
r	Ę	SEO	_		X	31		32		33		34		35	·	36		37		38		39		40		4	
			<u></u>		Vector	Uni-ZAP XR		Uni-ZAP XR		Uni-ZAP XR		pCMVSport	2.0	Uni-ZAP XR		Uni-ZAP XR		Uni-ZAP XR		Uni-ZAP XR		pBluescript		Uni-ZAP XR		Uni-ZAP XR	
		ATCC	Denosit	Nr and	Date	209628	02/12/98	209628	02/17/98	209628	02/12/98	209628	02/12/98	209628	02/12/98	209628	02/17/98	209628	02/17/98	209628	02/12/98	209628	02/17/98	209628	02/17/98	209628	02/17/28
				CDNA	Clone ID	HODCU34		HODCZ09		HOEDB32		HOGAG15		HPIBO48		HPMFP40		HPRCU95		HPTTG19		HPTVX32		HRDDV47		HRDEN56	
				Gene	Š.	21		22		23		24		25		26		27		28	l	29		30		31	

	_										_								- T				-		-		\neg
•		Last	AA	of	SS	2		62		8 8	:	45		52	j	9/	!	45		<u> </u>	-	133		49	1	63	
		First	AA of	Secreted	Portion	37	Š	20		56		34		38]	36		61		46		23		20		25	
	Last	AA	Jo		Pep	36	!	61		25	,	33		37		35		<u>∞</u>		45		77		61		24	
	First	AA	Jo	Sig	Pep					_		_		_		-				_		_		<u>.</u>			
	AA	SEQ	Ω	ÖN	>	139		140		141		142		143		144		145		146		147		148		149	
5' NT	Jo	First	AA of	Signal NO:	Pep	39		130		63		295		37		53		691		73		133		32		120	
		5' NT	Jo	Start	Codon	39		130		63		295		37		29		169		73		133		32		120	
	3, NT	Jo	Clone	Seq.		1162		654		1155		1112		4023		542		1495		818		1711		749		804	
	5' NT 3' NT	Jo	Clone Clone	Seq.				_		_						14		_		_				_		27	
			Total	Z	Seq.	1162		657		1155		1112		4023		542		1495		818		1711		749		1001	
Г	Z	SEO	_ □	Ö.	X	42		43		44		45		46		47		48		49		20		51		52	
					Vector	Uni-ZAP XR		Uni-ZAP XR		Uni-ZAP XR		Uni-ZAP XR		pSport1		Uni-ZAP XR		Uni-ZAP XR		Uni-ZAP XR		Uni-ZAP XR		Lambda ZAP	II	Uni-ZAP XR	
		ATCC	Deposit	Nr and	Date	209641	02/25/98	209641	02/25/98	209641	02/25/98	209641	02/25/98	209641	02/25/98	209641	02/25/98	209641	02/25/98	209641	02/25/98	209641	02/25/98	209641	02/25/98	209641	02/25/98
			-	cDNA	Clone ID	HSFAN12		HSQCM10		HSVAT68		HSXEC75		HTDAI54		HTEIT45		HTGBE48		HTLEP53		HTTBI76		HTWKG71		HTXDN32	
				Gene	So.	32		33		34		35		36		37		38		39		40		41		42	

-			ĹŢ.			Г		_		_		<u> </u>		Ė				Γ-	_	_				Γ	_
- 5	AA A		ORF	74		54		57		52		20		99		48		61		117		150		16	
	AA of	Secreted	Portion	21		33		41		16		30		34		24		56		21		23		30	
Last	of Of	Sig	Pep	20		32		40		15		29		33		23		25		20		22		56	
	of	Sig	Pep	1		_		1		-		_				_		1		-		1		1	
) []	NO:	Y	150		151		152		153		154		155		156		157		158		159		160	
5' NT of AA	~	-	Pep	61		125		19		149		123		56		170		132		85		306		09	
s, NT	Jo c	Start	Codon	61		125		19	:	149		123		97		170		132		85		306		09	
3, NT	Clone	Seq.		2254		486		1270		2059		898		986		695		314		734		1410		1231	
5, NT 3, NT	Total Clone Clone	Seq.		1		55		1		1		1		-		1		1		-		33		1	
		N	Seq.	2254		486		1270		502		898		986		695		314		734		1410		1231	
NT	Ω Ω	NO:	X	53		54		55		99		57		58		59		09		19		62		63	
			Vector	pBluescript		Uni-ZAP XR		Uni-ZAP XR		Uni-ZAP XR		pSport1		pCMVSport	3.0	Uni-ZAP XR									
ΔΤΟ	Deposit	Nr and	Date	209641	02/25/98	209641	02/25/98	209641	02/25/98	209641	02/25/98	209641	02/25/98	209641	02/25/98	209641	02/25/98	209641	02/25/98	209641	02/25/98	209641	02/25/98	209641	02/25/98
			Clone ID	HTSGX80		HTXEY51		HTXFH55		HTXJW17		HUFCJ30		HWAAP70		HWABW49		HWBDP28		HWDAC39		HWHGQ49		HJPAD75	
		Gene	Š Ž	43		44		45		46		47		48		49		20		51		52		53	

		Last	ΑA	Jo	ORF	41		77		54		68		44		44		65		61		46		105		44	
		First	AA of	Secreted	Portion	27		28		16		20		21		25		30		28		25		22		42	
	Last	AA		Sig		56		27		15		61		20		24		56		27		24		21		41	
	First	AA	Jo	Sig	Pep	1		1		ī		1		1		1		1		1		-		1		1	
Γ	AA	SEQ	Ω	NO:	Y	161		162		163		164		165		166		167		168		169		170		171	
5'NT	Jo	First	AA of	Signal NO:	Pep	215	. '	111		132		142		48		278		001		27		343		233		88	
		5' NT	Jo	Start	Codon	215		111		132		142		48		278		100		27		343		233		88	
	3, NT	Jo	Clone	Seq.		612		2231		1283		1123		1617		1389		9681		308		1688		1138		777	
	5' NT 3' NT	Jo	Clone Clone	Seq.		1		-		-		_		-		1						_		-		_	
			Total	Ľ	Seq.	612		2270		1283		1263		191		1389		1896		308		1688		1138		LLL	
	Z	SEQ	Ω	NO:	X	64		65		99		<i>L</i> 9		89		69		70		71		72		73		74	
					Vector	pCMVSport	3.0	Uni-ZAP XR		Uni-ZAP XR		pSport1		pCMVSport	3.0	Uni-ZAP XR		pBluescript	SK-	pBluescript		pSport1		pCMVSport	3.0	pCMVSport	3.0
		ATCC	Deposit	Nr and	Date	209641	02/25/98	209641	02/25/98	209641	02/25/98	209641	02/25/98	209641	02/25/98	209641	02/25/98	209641	02/25/98	209641	02/25/98	209641	02/25/98	209641	02/25/98	209641	02/25/98
				cDNA		HLDRP33		59МГЗМН		HNGFE55		HNKAA41		HRAAJ19		HSAWV96		HSBBT37		HSDZR57		81LISNH		HWBCP79		HYAAL70	
				Gene	No.	54		55		99		57		28		59		09		61		62		63		64	

	-			Ĺ	т			т	Τ'	1	1_	T	T		1	T
	Last	AA	of	ORF	4		4	46	40	52	249	147	47	55	48	45
	First	AA of	Secreted	Portion	27		29	35	22	29	27	31	16	15	25	26
Last	AA	Jo	Sig	Pep	26		28	34	21	28	26	30	15	14	24	25
First	AA	of	Sig	Pep	-		-	_	1	-	-	1	_	_	-	_
AA	SEQ	Ω	:0 	⊁	172		173	174	175	176	177	178	179	180	181	182
5° NT of	First SEQ	AA of	Signal NO:	Pep	118		252	87	06	10	104	180	142	334	309	206
	5°NT	of	Start	Codon	118		252	87	96	10	104	180	142	334	309	206
3, NT	of	Clone	Seq.		1060		1479	872	573	1509	1109	807	1043	1173	1561	1433
5' NT 3' NT	Jo	Clone Clone	Seq.		1		45	-	-	-		-	-	_	-	170
		Total	Z	Seq.	1060		1503	872	573	1509	1109	807	1043	1173	1561	1433
NT	SEQ	Ω	: ÖN	X	75		92	11	78	79	8	81	82	83	84	85
				Vector	pCMVSport	3.0	Uni-ZAP XR	Uni-ZAP XR	pSport1	Uni-ZAP XR	pSport1	pSport1	Uni-ZAP XR	Uni-ZAP XR	Uni-ZAP XR	Uni-ZAP XR
_	ATCC	Deposit	Nr and	Date	209641	02/25/98	209651	209651 03/04/98	209651 03/04/98	209651						
			cDNA	Clone ID	HYAAY86		HAPBS03	HBJLC01	HBLKD56	HCENK38	HCHMX01	HCHNF25	HE6GA29	HE6GE84	нетно95	HFCFJ18
			Gene	No.	65		99	29	89	69	70	71	72	73	74	75

	Last	AA	Jo	ORF	40		6/		46		140		338	Ì	65		360		191		159		100		42	
		AA of		Portion	36		16		27		25		76		27		36		7		15		25		27	
Last	AA	Jo	Sig	Pep	35		28		76		24		25		56		35		-		14		24		26	
First	AA	of	Sig	Pep	_				-		_						-		_		_		-		-	
AA	SEQ		ÖN	Υ	183		184		185		186		187		188		189		204		190		191		192	
5° NT of	First SEQ	AA of	Signal NO:	Pep	144		204		30		326		38		230		57		172		35		104		177	
	5° NT	Jo	Start	Codon	144		204		30		326		38		230		57		172		35		104		177	
3, NT		Clone	Seq.		1377		1715		417		1167		1892		523		1296		815		1747		009		286	
5' NT 3' NT	Jo	Total Clone Clone	Seq.		-1		_	-			304		_		57		630		1		1		1		1	
		Total	N	Seq.	1377		1715		417		1167		1892		523		1382		1201		1747		009		985	
Į.	SEQ	Ð	ö	X	98		87		88		68		06		91		92		107		63		94		95	
				Vector	Uni-ZAP XR		Lambda ZAP	II	pCMVSport 1		pCMVSport	3.0	pCMVSport	3.0	Uni-ZAP XR		Uni-ZAP XR		Uni-ZAP XR		pCMVSport	3.0	Uni-ZAP XR		Uni-ZAP XR	
	ATCC	Deposit	Nr and	Date	209651	03/04/98	209651	03/04/98	209651	03/04/98	209651	03/04/98	209651	03/04/98	209651	03/04/98	209651	03/04/98	209651	03/04/98	209651	03/04/98	209651	03/04/98	209651	03/04/98
			cDNA	Clone ID	HFPBM30		HFXKT05	,	HKB1E57		HLWAD77		HLWAY54		HNGBU28		НООНН51		нопнизі		HRAAB15		HSAVH65		HSDGN55	
			Gene	Š.	92		11		78		79		80		81		82		82		83		84		85	

									S' NT					Γ
				NT		5' NT 3' NT	3, NT		of	AA	First	Last		
		ATCC		SEQ		Jo	Jo	5° NT	First SEQ	SEQ	AA	AA	First	Last
		Deposit		Ω	Total	Clone Clone	Clone	Jo	AA of	Ð	Jo		AA of	AA
Gene	cDNA	Nr and		Ö	TN	Seq.	Sed.	Start	Signal NO:	: 	Sig		Secreted	Jo
No.	Clone ID	Date	Vector	×	Seq.			Codon	Pep	Χ	Pep	Pep	Portion	ORF
98	HSXAH81	209651	Uni-ZAP XR	96	802	1	802	88	88	193		21	22	19
		03/04/98												
87	HSXBX80	209651	Uni-ZAP XR	26	1226	-	1226	77	11	194		22	23	22
		03/04/98												
88	HTEHV08	209651	Uni-ZAP XR	86	1120	1	1120	382	382	195	_	17	18	185
		03/04/98												
68	HUFAK67	209651	pSport1	66	2596	1	2596	225	225	196	_	21	77	9/
		03/04/98												
06	HUSXS50	209651	pSport1	100	1020	_	1020	179	179	161	_	23	24	174
		03/04/98	i											
16	HAPON17	209651	Uni-ZAP XR	101	1520		1520	766	566	198	_	23	24	20
		03/04/98												
92	HATAC53	209651	Uni-ZAP XR	102	1306	13	1306	66	66	199		21	22	189
		03/04/98												
93	HAMFK58	209641	pCMVSport	103	785	-	785	279	279	200		3	32	79
		02/25/98	3.0											
94	HLYCH68	209641	pSport1	104	2015	34	1571	81	8	701	_	19	20	105
		02/25/98												
95	HCUHK65	209641	ZAP Express	105	367		367	08	8	202	_	26	27	79
		02/25/98												

15

20

25

30

Table 1 summarizes the information corresponding to each "Gene No." described above. The nucleotide sequence identified as "NT SEQ ID NO:X" was assembled from partially homologous ("overlapping") sequences obtained from the "cDNA clone ID" identified in Table 1 and, in some cases, from additional related DNA clones. The overlapping sequences were assembled into a single contiguous sequence of high redundancy (usually three to five overlapping sequences at each nucleotide position), resulting in a final sequence identified as SEQ ID NO:X.

The cDNA Clone ID was deposited on the date and given the corresponding deposit number listed in "ATCC Deposit No:Z and Date." Some of the deposits contain multiple different clones corresponding to the same gene. "Vector" refers to the type of vector contained in the cDNA Clone ID.

"Total NT Seq." refers to the total number of nucleotides in the contig identified by "Gene No." The deposited clone may contain all or most of these sequences, reflected by the nucleotide position indicated as "5' NT of Clone Seq." and the "3' NT of Clone Seq." of SEQ ID NO:X. The nucleotide position of SEQ ID NO:X of the putative start codon (methionine) is identified as "5' NT of Start Codon." Similarly, the nucleotide position of SEQ ID NO:X of the predicted signal sequence is identified as "5' NT of First AA of Signal Pep."

The translated amino acid sequence, beginning with the methionine, is identified as "AA SEQ ID NO:Y," although other reading frames can also be easily translated using known molecular biology techniques. The polypeptides produced by these alternative open reading frames are specifically contemplated by the present invention.

The first and last amino acid position of SEQ ID NO:Y of the predicted signal peptide is identified as "First AA of Sig Pep" and "Last AA of Sig Pep." The predicted first amino acid position of SEQ ID NO:Y of the secreted portion is identified as "Predicted First AA of Secreted Portion." Finally, the amino acid position of SEQ ID NO:Y of the last amino acid in the open reading frame is identified as "Last AA of ORF."

SEQ ID NO:X and the translated SEQ ID NO:Y are sufficiently accurate and otherwise suitable for a variety of uses well known in the art and described further

190

below. For instance, SEQ ID NO:X is useful for designing nucleic acid hybridization probes that will detect nucleic acid sequences contained in SEQ ID NO:X or the cDNA contained in the deposited clone. These probes will also hybridize to nucleic acid molecules in biological samples, thereby enabling a variety of forensic and diagnostic methods of the invention. Similarly, polypeptides identified from SEQ ID NO:Y may be used to generate antibodies which bind specifically to the secreted proteins encoded by the cDNA clones identified in Table 1.

5

10

15

20

25

30

Nevertheless, DNA sequences generated by sequencing reactions can contain sequencing errors. The errors exist as misidentified nucleotides, or as insertions or deletions of nucleotides in the generated DNA sequence. The erroneously inserted or deleted nucleotides cause frame shifts in the reading frames of the predicted amino acid sequence. In these cases, the predicted amino acid sequence diverges from the actual amino acid sequence, even though the generated DNA sequence may be greater than 99.9% identical to the actual DNA sequence (for example, one base insertion or deletion in an open reading frame of over 1000 bases).

Accordingly, for those applications requiring precision in the nucleotide sequence or the amino acid sequence, the present invention provides not only the generated nucleotide sequence identified as SEQ ID NO:X and the predicted translated amino acid sequence identified as SEQ ID NO:Y, but also a sample of plasmid DNA containing a human cDNA of the invention deposited with the ATCC, as set forth in Table 1. The nucleotide sequence of each deposited clone can readily be determined by sequencing the deposited clone in accordance with known methods. The predicted amino acid sequence can then be verified from such deposits.

Moreover, the amino acid sequence of the protein encoded by a particular clone can also be directly determined by peptide sequencing or by expressing the protein in a suitable host cell containing the deposited human cDNA, collecting the protein, and determining its sequence.

The present invention also relates to the genes corresponding to SEQ ID NO:X, SEQ ID NO:Y, or the deposited clone. The corresponding gene can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include preparing probes or primers from the disclosed

WO 99/47540

191

PCT/US99/05804

sequence and identifying or amplifying the corresponding gene from appropriate sources of genomic material.

Also provided in the present invention are species homologs. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source for the desired homologue.

The polypeptides of the invention can be prepared in any suitable manner.

Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

The polypeptides may be in the form of the secreted protein, including the mature form, or may be a part of a larger protein, such as a fusion protein (see below). It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences which aid in purification, such as multiple histidine residues, or an additional sequence for stability during recombinant production.

The polypeptides of the present invention are preferably provided in an isolated form, and preferably are substantially purified. A recombinantly produced version of a polypeptide, including the secreted polypeptide, can be substantially purified by the one-step method described in Smith and Johnson, Gene 67:31-40 (1988). Polypeptides of the invention also can be purified from natural or recombinant sources using antibodies of the invention raised against the secreted protein in methods which are well known in the art.

25

30

5

10

15

20

Signal Sequences

Methods for predicting whether a protein has a signal sequence, as well as the cleavage point for that sequence, are available. For instance, the method of McGeoch, Virus Res. 3:271-286 (1985), uses the information from a short N-terminal charged region and a subsequent uncharged region of the complete (uncleaved) protein. The method of von Heinje, Nucleic Acids Res. 14:4683-4690 (1986) uses the

information from the residues surrounding the cleavage site, typically residues -13 to +2, where +1 indicates the amino terminus of the secreted protein. The accuracy of predicting the cleavage points of known mammalian secretory proteins for each of these methods is in the range of 75-80%. (von Heinje, supra.) However, the two methods do not always produce the same predicted cleavage point(s) for a given protein.

In the present case, the deduced amino acid sequence of the secreted polypeptide was analyzed by a computer program called SignalP (Henrik Nielsen et al., Protein Engineering 10:1-6 (1997)), which predicts the cellular location of a protein based on the amino acid sequence. As part of this computational prediction of localization, the methods of McGeoch and von Heinje are incorporated. The analysis of the amino acid sequences of the secreted proteins described herein by this program provided the results shown in Table 1.

As one of ordinary skill would appreciate, however, cleavage sites sometimes vary from organism to organism and cannot be predicted with absolute certainty. Accordingly, the present invention provides secreted polypeptides having a sequence shown in SEQ ID NO:Y which have an N-terminus beginning within 5 residues (i.e., + or - 5 residues) of the predicted cleavage point. Similarly, it is also recognized that in some cases, cleavage of the signal sequence from a secreted protein is not entirely uniform, resulting in more than one secreted species. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

Moreover, the signal sequence identified by the above analysis may not necessarily predict the naturally occurring signal sequence. For example, the naturally occurring signal sequence may be further upstream from the predicted signal sequence. However, it is likely that the predicted signal sequence will be capable of directing the secreted protein to the ER. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

Polynucleotide and Polypeptide Variants

5

10

15

20

25

30

10

15

20

25

30

PCT/US99/05804

"Variant" refers to a polynucleotide or polypeptide differing from the polynucleotide or polypeptide of the present invention, but retaining essential properties thereof. Generally, variants are overall closely similar, and, in many regions, identical to the polynucleotide or polypeptide of the present invention.

By a polynucleotide having a nucleotide sequence at least, for example, 95% "identical" to a reference nucleotide sequence of the present invention, it is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence encoding the polypeptide. In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. The query sequence may be an entire sequence shown in Table 1, the ORF (open reading frame), or any fragement specified as described herein.

As a practical matter, whether any particular nucleic acid molecule or polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to a nucleotide sequence of the presence invention can be determined conventionally using known computer programs. A preferred method for determing the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245). In a sequence alignment the query and subject sequences are both DNA sequences. An RNA sequence can be compared by converting U's to T's. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB alignment of DNA sequences to calculate percent identity are: Matrix=Unitary, k-tuple=4, Mismatch Penalty=1, Joining Penalty=30, Randomization Group Length=0, Cutoff Score=1, Gap Penalty=5, Gap Size Penalty 0.05, Window Size=500 or the length of the subject nucleotide sequence, whichever is shorter.

10

15

20

25

30

If the subject sequence is shorter than the query sequence because of 5' or 3' deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for 5' and 3' truncations of the subject sequence when calculating percent identity. For subject sequences truncated at the 5' or 3' ends, relative to the the query sequence, the percent identity is corrected by calculating the number of bases of the query sequence that are 5' and 3' of the subject sequence, which are not matched/aligned, as a percent of the total bases of the query sequence. Whether a nucleotide is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This corrected score is what is used for the purposes of the present invention. Only bases outside the 5' and 3' bases of the subject sequence, as displayed by the FASTDB alignment, which are not matched/aligned with the query sequence, are calculated for the purposes of manually adjusting the percent identity score.

For example, a 90 base subject sequence is aligned to a 100 base query sequence to determine percent identity. The deletions occur at the 5' end of the subject sequence and therefore, the FASTDB alignment does not show a matched/alignement of the first 10 bases at 5' end. The 10 unpaired bases represent 10% of the sequence (number of bases at the 5' and 3' ends not matched/total number of bases in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 bases were perfectly matched the final percent identity would be 90%. In another example, a 90 base subject sequence is compared with a 100 base query sequence. This time the deletions are internal deletions so that there are no bases on the 5' or 3' of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only bases 5' and 3' of the subject sequence which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are to made for the purposes of the present invention.

By a polypeptide having an amino acid sequence at least, for example, 95%

195

"identical" to a query amino acid sequence of the present invention, it is intended that the amino acid sequence of the subject polypeptide is identical to the query sequence except that the subject polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the guery amino acid sequence. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a query amino acid sequence, up to 5% of the amino acid residues in the subject sequence may be inserted, deleted, (indels) or substituted with another amino acid. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

5

10

15

25

30

As a practical matter, whether any particular polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to, for instance, the amino acid sequences shown in Table 1 or to the amino acid sequence encoded by deposited DNA clone can be determined conventionally using known computer programs. A preferred method for determing the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245). In a sequence alignment the 20 query and subject sequences are either both nucleotide sequences or both amino acid sequences. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB amino acid alignment are: Matrix=PAM 0, k-tuple=2, Mismatch Penalty=1, Joining Penalty=20, Randomization Group Length=0, Cutoff Score=1, Window Size=sequence length, Gap Penalty=5, Gap Size Penalty=0.05, Window Size=500 or the length of the subject amino acid sequence, whichever is shorter.

If the subject sequence is shorter than the query sequence due to N- or Cterminal deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for Nand C-terminal truncations of the subject sequence when calculating global percent identity. For subject sequences truncated at the N- and C-termini, relative to the the

196

query sequence, the percent identity is corrected by calculating the number of residues of the query sequence that are N- and C-terminal of the subject sequence, which are not matched/aligned with a corresponding subject residue, as a percent of the total bases of the query sequence. Whether a residue is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This final percent identity score is what is used for the purposes of the present invention. Only residues to the N- and C-termini of the subject sequence, which are not matched/aligned with the query sequence, are considered for the purposes of manually adjusting the percent identity score. That is, only query residue positions outside the farthest N- and C-terminal residues of the subject sequence.

5

10

15

20

25

30

For example, a 90 amino acid residue subject sequence is aligned with a 100 residue query sequence to determine percent identity. The deletion occurs at the Nterminus of the subject sequence and therefore, the FASTDB alignment does not show a matching/alignment of the first 10 residues at the N-terminus. The 10 unpaired residues represent 10% of the sequence (number of residues at the N- and Ctermini not matched/total number of residues in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 residues were perfectly matched the final percent identity would be 90%. In another example, a 90 residue subject sequence is compared with a 100 residue query sequence. This time the deletions are internal deletions so there are no residues at the N- or C-termini of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only residue positions outside the N- and C-terminal ends of the subject sequence, as displayed in the FASTDB alignment, which are not matched/aligned with the query sequnce are manually corrected for. No other manual corrections are to made for the purposes of the present invention.

The variants may contain alterations in the coding regions, non-coding regions, or both. Especially preferred are polynucleotide variants containing alterations which produce silent substitutions, additions, or deletions, but do not alter

10

15

20

PCT/US99/05804

the properties or activities of the encoded polypeptide. Nucleotide variants produced by silent substitutions due to the degeneracy of the genetic code are preferred. Moreover, variants in which 5-10, 1-5, or 1-2 amino acids are substituted, deleted, or added in any combination are also preferred. Polynucleotide variants can be produced for a variety of reasons, e.g., to optimize codon expression for a particular host (change codons in the human mRNA to those preferred by a bacterial host such as E. coli).

Naturally occurring variants are called "allelic variants," and refer to one of several alternate forms of a gene occupying a given locus on a chromosome of an organism. (Genes II, Lewin, B., ed., John Wiley & Sons, New York (1985).) These allelic variants can vary at either the polynucleotide and/or polypeptide level. Alternatively, non-naturally occurring variants may be produced by mutagenesis techniques or by direct synthesis.

Using known methods of protein engineering and recombinant DNA technology, variants may be generated to improve or alter the characteristics of the polypeptides of the present invention. For instance, one or more amino acids can be deleted from the N-terminus or C-terminus of the secreted protein without substantial loss of biological function. The authors of Ron et al., J. Biol. Chem. 268: 2984-2988 (1993), reported variant KGF proteins having heparin binding activity even after deleting 3, 8, or 27 amino-terminal amino acid residues. Similarly, Interferon gamma exhibited up to ten times higher activity after deleting 8-10 amino acid residues from the carboxy terminus of this protein. (Dobeli et al., J. Biotechnology 7:199-216 (1988).)

Moreover, ample evidence demonstrates that variants often retain a biological activity similar to that of the naturally occurring protein. For example, Gayle and coworkers (J. Biol. Chem 268:22105-22111 (1993)) conducted extensive mutational analysis of human cytokine IL-1a. They used random mutagenesis to generate over 3,500 individual IL-1a mutants that averaged 2.5 amino acid changes per variant over the entire length of the molecule. Multiple mutations were examined at every possible amino acid position. The investigators found that "[m]ost of the molecule could be altered with little effect on either [binding or biological activity]." (See,

10

15

20

25

30

PCT/US99/05804

Abstract.) In fact, only 23 unique amino acid sequences, out of more than 3,500 nucleotide sequences examined, produced a protein that significantly differed in activity from wild-type.

Furthermore, even if deleting one or more amino acids from the N-terminus or C-terminus of a polypeptide results in modification or loss of one or more biological functions, other biological activities may still be retained. For example, the ability of a deletion variant to induce and/or to bind antibodies which recognize the secreted form will likely be retained when less than the majority of the residues of the secreted form are removed from the N-terminus or C-terminus. Whether a particular polypeptide lacking N- or C-terminal residues of a protein retains such immunogenic activities can readily be determined by routine methods described herein and otherwise known in the art.

Thus, the invention further includes polypeptide variants which show substantial biological activity. Such variants include deletions, insertions, inversions, repeats, and substitutions selected according to general rules known in the art so as have little effect on activity. For example, guidance concerning how to make phenotypically silent amino acid substitutions is provided in Bowie, J. U. et al., Science 247:1306-1310 (1990), wherein the authors indicate that there are two main strategies for studying the tolerance of an amino acid sequence to change.

The first strategy exploits the tolerance of amino acid substitutions by natural selection during the process of evolution. By comparing amino acid sequences in different species, conserved amino acids can be identified. These conserved amino acids are likely important for protein function. In contrast, the amino acid positions where substitutions have been tolerated by natural selection indicates that these positions are not critical for protein function. Thus, positions tolerating amino acid substitution could be modified while still maintaining biological activity of the protein.

The second strategy uses genetic engineering to introduce amino acid changes at specific positions of a cloned gene to identify regions critical for protein function. For example, site directed mutagenesis or alanine-scanning mutagenesis (introduction of single alanine mutations at every residue in the molecule) can be used.

199

(Cunningham and Wells, Science 244:1081-1085 (1989).) The resulting mutant molecules can then be tested for biological activity.

5

10

15

20

25

30

As the authors state, these two strategies have revealed that proteins are surprisingly tolerant of amino acid substitutions. The authors further indicate which amino acid changes are likely to be permissive at certain amino acid positions in the protein. For example, most buried (within the tertiary structure of the protein) amino acid residues require nonpolar side chains, whereas few features of surface side chains are generally conserved. Moreover, tolerated conservative amino acid substitutions involve replacement of the aliphatic or hydrophobic amino acids Ala, Val, Leu and Ile; replacement of the hydroxyl residues Ser and Thr; replacement of the acidic residues Asp and Glu; replacement of the amide residues Asn and Gln, replacement of the basic residues Lys, Arg, and His; replacement of the aromatic residues Phe, Tyr, and Trp, and replacement of the small-sized amino acids Ala, Ser, Thr, Met, and Gly.

Besides conservative amino acid substitution, variants of the present invention include (i) substitutions with one or more of the non-conserved amino acid residues, where the substituted amino acid residues may or may not be one encoded by the genetic code, or (ii) substitution with one or more of amino acid residues having a substituent group, or (iii) fusion of the mature polypeptide with another compound, such as a compound to increase the stability and/or solubility of the polypeptide (for example, polyethylene glycol), or (iv) fusion of the polypeptide with additional amino acids, such as an IgG Fc fusion region peptide, or leader or secretory sequence, or a sequence facilitating purification. Such variant polypeptides are deemed to be within the scope of those skilled in the art from the teachings herein.

For example, polypeptide variants containing amino acid substitutions of charged amino acids with other charged or neutral amino acids may produce proteins with improved characteristics, such as less aggregation. Aggregation of pharmaceutical formulations both reduces activity and increases clearance due to the aggregate's immunogenic activity. (Pinckard et al., Clin. Exp. Immunol. 2:331-340 (1967); Robbins et al., Diabetes 36: 838-845 (1987); Cleland et al., Crit. Rev. Therapeutic Drug Carrier Systems 10:307-377 (1993).)

WO 99/47540

A further embodiment of the invention relates to a polypeptide which comprises the amino acid sequence of the present invention having an amino acid sequence which contains at least one amino acid substitution, but not more than 50 amino acid substitutions, even more preferably, not more than 40 amino acid substitutions, still more preferably, not more than 30 amino acid substitutions, and still even more preferably, not more than 20 amino acid substitutions. Of course, in order of ever-increasing preference, it is highly preferable for a polypeptide to have an amino acid sequence which comprises the amino acid sequence of the present invention, which contains at least one, but not more than 10, 9, 8, 7, 6, 5, 4, 3, 2 or 1 amino acid substitutions. In specific embodiments, the number of additions, substitutions, and/or deletions in the amino acid sequence of the present invention or fragments thereof (e.g., the mature form and/or other fragments described herein), is 1-5, 5-10, 5-25, 5-50, 10-50 or 50-150, conservative amino acid substitutions are preferable.

15

20

25

30

10

5

Polynucleotide and Polypeptide Fragments

In the present invention, a "polynucleotide fragment" refers to a short polynucleotide having a nucleic acid sequence contained in the deposited clone or shown in SEQ ID NO:X. The short nucleotide fragments are preferably at least about 15 nt, and more preferably at least about 20 nt, still more preferably at least about 30 nt, and even more preferably, at least about 40 nt in length. A fragment "at least 20 nt in length," for example, is intended to include 20 or more contiguous bases from the cDNA sequence contained in the deposited clone or the nucleotide sequence shown in SEQ ID NO:X. These nucleotide fragments are useful as diagnostic probes and primers as discussed herein. Of course, larger fragments (e.g., 50, 150, 500, 600, 2000 nucleotides) are preferred.

Moreover, representative examples of polynucleotide fragments of the invention, include, for example, fragments having a sequence from about nucleotide number 1-50, 51-100, 101-150, 151-200, 201-250, 251-300, 301-350, 351-400, 401-450, 451-500, 501-550, 551-600, 651-700, 701-750, 751-800, 800-850, 851-900, 901-950, 951-1000, 1001-1050, 1051-1100, 1101-1150, 1151-1200, 1201-1250, 1251-

10

15

20

25

30

1300, 1301-1350, 1351-1400, 1401-1450, 1451-1500, 1501-1550, 1551-1600, 1601-1650, 1651-1700, 1701-1750, 1751-1800, 1801-1850, 1851-1900, 1901-1950, 1951-2000, or 2001 to the end of SEQ ID NO:X or the cDNA contained in the deposited clone. In this context "about" includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) nucleotides, at either terminus or at both termini. Preferably, these fragments encode a polypeptide which has biological activity. More preferably, these polynucleotides can be used as probes or primers as discussed herein.

In the present invention, a "polypeptide fragment" refers to a short amino acid sequence contained in SEQ ID NO:Y or encoded by the cDNA contained in the deposited clone. Protein fragments may be "free-standing," or comprised within a larger polypeptide of which the fragment forms a part or region, most preferably as a single continuous region. Representative examples of polypeptide fragments of the invention, include, for example, fragments from about amino acid number 1-20, 21-40, 41-60, 61-80, 81-100, 102-120, 121-140, 141-160, or 161 to the end of the coding region. Moreover, polypeptide fragments can be about 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, or 150 amino acids in length. In this context "about" includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) amino acids, at either extreme or at both extremes.

Preferred polypeptide fragments include the secreted protein as well as the mature form. Further preferred polypeptide fragments include the secreted protein or the mature form having a continuous series of deleted residues from the amino or the carboxy terminus, or both. For example, any number of amino acids, ranging from 1-60, can be deleted from the amino terminus of either the secreted polypeptide or the mature form. Similarly, any number of amino acids, ranging from 1-30, can be deleted from the carboxy terminus of the secreted protein or mature form. Furthermore, any combination of the above amino and carboxy terminus deletions are preferred. Similarly, polynucleotide fragments encoding these polypeptide fragments are also preferred.

Also preferred are polypeptide and polynucleotide fragments characterized by structural or functional domains, such as fragments that comprise alpha-helix and

10

15

20

25

30

alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions. Polypeptide fragments of SEQ ID NO:Y falling within conserved domains are specifically contemplated by the present invention. Moreover, polynucleotide fragments encoding these domains are also contemplated.

Other preferred fragments are biologically active fragments. Biologically active fragments are those exhibiting activity similar, but not necessarily identical, to an activity of the polypeptide of the present invention. The biological activity of the fragments may include an improved desired activity, or a decreased undesirable activity.

Epitopes & Antibodies

In the present invention, "epitopes" refer to polypeptide fragments having antigenic or immunogenic activity in an animal, especially in a human. A preferred embodiment of the present invention relates to a polypeptide fragment comprising an epitope, as well as the polynucleotide encoding this fragment. A region of a protein molecule to which an antibody can bind is defined as an "antigenic epitope." In contrast, an "immunogenic epitope" is defined as a part of a protein that elicits an antibody response. (See, for instance, Geysen et al., Proc. Natl. Acad. Sci. USA 81:3998- 4002 (1983).)

Fragments which function as epitopes may be produced by any conventional means. (See, e.g., Houghten, R. A., Proc. Natl. Acad. Sci. USA 82:5131-5135 (1985) further described in U.S. Patent No. 4,631,211.)

In the present invention, antigenic epitopes preferably contain a sequence of at least seven, more preferably at least nine, and most preferably between about 15 to about 30 amino acids. Antigenic epitopes are useful to raise antibodies, including monoclonal antibodies, that specifically bind the epitope. (See, for instance, Wilson et al., Cell 37:767-778 (1984); Sutcliffe, J. G. et al., Science 219:660-666 (1983).)

WO 99/47540

203

PCT/US99/05804

Similarly, immunogenic epitopes can be used to induce antibodies according to methods well known in the art. (See, for instance, Sutcliffe et al., supra; Wilson et al., supra; Chow, M. et al., Proc. Natl. Acad. Sci. USA 82:910-914; and Bittle, F. J. et al., J. Gen. Virol. 66:2347-2354 (1985).) A preferred immunogenic epitope includes the secreted protein. The immunogenic epitopes may be presented together with a carrier protein, such as an albumin, to an animal system (such as rabbit or mouse) or, if it is long enough (at least about 25 amino acids), without a carrier. However, immunogenic epitopes comprising as few as 8 to 10 amino acids have been shown to be sufficient to raise antibodies capable of binding to, at the very least, linear epitopes in a denatured polypeptide (e.g., in Western blotting.)

As used herein, the term "antibody" (Ab) or "monoclonal antibody" (Mab) is meant to include intact molecules as well as antibody fragments (such as, for example, Fab and F(ab')2 fragments) which are capable of specifically binding to protein. Fab and F(ab')2 fragments lack the Fc fragment of intact antibody, clear more rapidly from the circulation, and may have less non-specific tissue binding than an intact antibody. (Wahl et al., J. Nucl. Med. 24:316-325 (1983).) Thus, these fragments are preferred, as well as the products of a FAB or other immunoglobulin expression library. Moreover, antibodies of the present invention include chimeric, single chain, and humanized antibodies.

20

25

30

5

10

15

Fusion Proteins

Any polypeptide of the present invention can be used to generate fusion proteins. For example, the polypeptide of the present invention, when fused to a second protein, can be used as an antigenic tag. Antibodies raised against the polypeptide of the present invention can be used to indirectly detect the second protein by binding to the polypeptide. Moreover, because secreted proteins target cellular locations based on trafficking signals, the polypeptides of the present invention can be used as targeting molecules once fused to other proteins.

Examples of domains that can be fused to polypeptides of the present invention include not only heterologous signal sequences, but also other heterologous

204

functional regions. The fusion does not necessarily need to be direct, but may occur through linker sequences.

Moreover, fusion proteins may also be engineered to improve characteristics of the polypeptide of the present invention. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence during purification from the host cell or subsequent handling and storage. Also, peptide moieties may be added to the polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to facilitate handling of polypeptides are familiar and routine techniques in the art.

10

15

20

25

30

Moreover, polypeptides of the present invention, including fragments, and specifically epitopes, can be combined with parts of the constant domain of immunoglobulins (IgG), resulting in chimeric polypeptides. These fusion proteins facilitate purification and show an increased half-life in vivo. One reported example describes chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. (EP A 394,827; Traunecker et al., Nature 331:84-86 (1988).) Fusion proteins having disulfide-linked dimeric structures (due to the IgG) can also be more efficient in binding and neutralizing other molecules, than the monomeric secreted protein or protein fragment alone. (Fountoulakis et al., J. Biochem. 270:3958-3964 (1995).)

Similarly, EP-A-O 464 533 (Canadian counterpart 2045869) discloses fusion proteins comprising various portions of constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, the Fc part in a fusion protein is beneficial in therapy and diagnosis, and thus can result in, for example, improved pharmacokinetic properties. (EP-A 0232 262.) Alternatively, deleting the Fc part after the fusion protein has been expressed, detected, and purified, would be desired. For example, the Fc portion may hinder therapy and diagnosis if the fusion protein is used as an antigen for immunizations. In drug discovery, for example, human proteins, such as hIL-5, have been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists of hIL-5. (See,

D. Bennett et al., J. Molecular Recognition 8:52-58 (1995); K. Johanson et al., J. Biol. Chem. 270:9459-9471 (1995).)

Moreover, the polypeptides of the present invention can be fused to marker sequences, such as a peptide which facilitates purification of the fused polypeptide.

In preferred embodiments, the marker amino acid sequence is a hexa-histidine peptide, such as the tag provided in a pQE vector (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311), among others, many of which are commercially available. As described in Gentz et al., Proc. Natl. Acad. Sci. USA 86:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein.

Another peptide tag useful for purification, the "HA" tag, corresponds to an epitope derived from the influenza hemagglutinin protein. (Wilson et al., Cell 37:767 (1984).)

Thus, any of these above fusions can be engineered using the polynucleotides or the polypeptides of the present invention.

15

20

25

30

Vectors, Host Cells, and Protein Production

The present invention also relates to vectors containing the polynucleotide of the present invention, host cells, and the production of polypeptides by recombinant techniques. The vector may be, for example, a phage, plasmid, viral, or retroviral vector. Retroviral vectors may be replication competent or replication defective. In the latter case, viral propagation generally will occur only in complementing host cells.

The polynucleotides may be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged in vitro using an appropriate packaging cell line and then transduced into host cells.

The polynucleotide insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the E. coli lac, trp, phoA and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The

206

expression constructs will further contain sites for transcription initiation, termination, and, in the transcribed region, a ribosome binding site for translation. The coding portion of the transcripts expressed by the constructs will preferably include a translation initiating codon at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the polypeptide to be translated.

As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase, G418 or neomycin resistance for eukaryotic cell culture and tetracycline, kanamycin or ampicillin resistance genes for culturing in E. coli and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as E. coli, Streptomyces and Salmonella typhimurium cells; fungal cells, such as yeast cells; insect cells such as Drosophila S2 and Spodoptera Sf9 cells; animal cells such as CHO, COS, 293, and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

10

15

20

25

30

Among vectors preferred for use in bacteria include pQE70, pQE60 and pQE-9, available from QIAGEN, Inc.; pBluescript vectors, Phagescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene Cloning Systems, Inc.; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia Biotech, Inc. Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Other suitable vectors will be readily apparent to the skilled artisan.

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection, or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al., Basic Methods In Molecular Biology (1986). It is specifically contemplated that the polypeptides of the present invention may in fact be expressed by a host cell lacking a recombinant vector.

A polypeptide of this invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography,

phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification.

Polypeptides of the present invention, and preferably the secreted form, can also be recovered from: products purified from natural sources, including bodily fluids, tissues and cells, whether directly isolated or cultured; products of chemical synthetic procedures; and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect, and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes. Thus, it is well known in the art that the N-terminal methionine encoded by the translation initiation codon generally is removed with high efficiency from any protein after translation in all eukaryotic cells. While the N-terminal methionine on most proteins also is efficiently removed in most prokaryotes, for some proteins, this prokaryotic removal process is inefficient, depending on the nature of the amino acid to which the N-terminal methionine is covalently linked.

In addition to encompassing host cells containing the vector constructs discussed herein, the invention also encompasses primary, secondary, and immortalized host cells of vertebrate origin, particularly mammalian origin, that have been engineered to delete or replace endogenous genetic material (e.g., coding sequence), and/or to include genetic material (e.g., heterologous polynucleotide sequences) that is operably associated with the polynucleotides of the invention, and which activates, alters, and/or amplifies endogenous polynucleotides. For example, techniques known in the art may be used to operably associate heterologous control regions (e.g., promoter and/or enhancer) and endogenous polynucleotide sequences via homologous recombination (see, e.g., U.S. Patent No. 5,641,670, issued June 24, 1997; International Publication No. WO 96/29411, published September 26, 1996; International Publication No. WO 94/12650, published August 4, 1994; Koller et al.,

Proc. Natl. Acad. Sci. USA 86:8932-8935 (1989); and Zijlstra et al., Nature 342:435-438 (1989), the disclosures of each of which are incorporated by reference in their entireties).

5

10

15

20

25

30

Uses of the Polynucleotides

Each of the polynucleotides identified herein can be used in numerous ways as reagents. The following description should be considered exemplary and utilizes known techniques.

The polynucleotides of the present invention are useful for chromosome identification. There exists an ongoing need to identify new chromosome markers, since few chromosome marking reagents, based on actual sequence data (repeat polymorphisms), are presently available. Each polynucleotide of the present invention can be used as a chromosome marker.

Briefly, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the sequences shown in SEQ ID NO:X. Primers can be selected using computer analysis so that primers do not span more than one predicted exon in the genomic DNA. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the SEQ ID NO:X will yield an amplified fragment.

Similarly, somatic hybrids provide a rapid method of PCR mapping the polynucleotides to particular chromosomes. Three or more clones can be assigned per day using a single thermal cycler. Moreover, sublocalization of the polynucleotides can be achieved with panels of specific chromosome fragments. Other gene mapping strategies that can be used include in situ hybridization, prescreening with labeled flow-sorted chromosomes, and preselection by hybridization to construct chromosome specific-cDNA libraries.

Precise chromosomal location of the polynucleotides can also be achieved using fluorescence in situ hybridization (FISH) of a metaphase chromosomal spread. This technique uses polynucleotides as short as 500 or 600 bases; however,

polynucleotides 2,000-4,000 bp are preferred. For a review of this technique, see Verma et al., "Human Chromosomes: a Manual of Basic Techniques," Pergamon Press, New York (1988).

5

10

15

20

25

30

For chromosome mapping, the polynucleotides can be used individually (to mark a single chromosome or a single site on that chromosome) or in panels (for marking multiple sites and/or multiple chromosomes). Preferred polynucleotides correspond to the noncoding regions of the cDNAs because the coding sequences are more likely conserved within gene families, thus increasing the chance of cross hybridization during chromosomal mapping.

Once a polynucleotide has been mapped to a precise chromosomal location, the physical position of the polynucleotide can be used in linkage analysis. Linkage analysis establishes coinheritance between a chromosomal location and presentation of a particular disease. (Disease mapping data are found, for example, in V. McKusick, Mendelian Inheritance in Man (available on line through Johns Hopkins University Welch Medical Library).) Assuming 1 megabase mapping resolution and one gene per 20 kb, a cDNA precisely localized to a chromosomal region associated with the disease could be one of 50-500 potential causative genes.

Thus, once coinheritance is established, differences in the polynucleotide and the corresponding gene between affected and unaffected individuals can be examined. First, visible structural alterations in the chromosomes, such as deletions or translocations, are examined in chromosome spreads or by PCR. If no structural alterations exist, the presence of point mutations are ascertained. Mutations observed in some or all affected individuals, but not in normal individuals, indicates that the mutation may cause the disease. However, complete sequencing of the polypeptide and the corresponding gene from several normal individuals is required to distinguish the mutation from a polymorphism. If a new polymorphism is identified, this polymorphic polypeptide can be used for further linkage analysis.

Furthermore, increased or decreased expression of the gene in affected individuals as compared to unaffected individuals can be assessed using polynucleotides of the present invention. Any of these alterations (altered expression,

PCT/US99/05804

chromosomal rearrangement, or mutation) can be used as a diagnostic or prognostic marker.

In addition to the foregoing, a polynucleotide can be used to control gene expression through triple helix formation or antisense DNA or RNA. Both methods rely on binding of the polynucleotide to DNA or RNA. For these techniques, 5 preferred polynucleotides are usually 20 to 40 bases in length and complementary to either the region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J. 10 Neurochem. 56:560 (1991); Oligodeoxy-nucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988).) Triple helix formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques are effective in model systems, and the information disclosed herein can be used to design antisense or triple helix polynucleotides in an effort to treat disease. 15

Polynucleotides of the present invention are also useful in gene therapy. One goal of gene therapy is to insert a normal gene into an organism having a defective gene, in an effort to correct the genetic defect. The polynucleotides disclosed in the present invention offer a means of targeting such genetic defects in a highly accurate manner. Another goal is to insert a new gene that was not present in the host genome, thereby producing a new trait in the host cell.

20

25

30

The polynucleotides are also useful for identifying individuals from minute biological samples. The United States military, for example, is considering the use of restriction fragment length polymorphism (RFLP) for identification of its personnel. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identifying personnel. This method does not suffer from the current limitations of "Dog Tags" which can be lost, switched, or stolen, making positive identification difficult. The polynucleotides of the present invention can be used as additional DNA markers for RFLP.

WO 99/47540

10

15

20

25

30

The polynucleotides of the present invention can also be used as an alternative to RFLP, by determining the actual base-by-base DNA sequence of selected portions of an individual's genome. These sequences can be used to prepare PCR primers for amplifying and isolating such selected DNA, which can then be sequenced. Using this technique, individuals can be identified because each individual will have a unique set of DNA sequences. Once an unique ID database is established for an individual, positive identification of that individual, living or dead, can be made from extremely small tissue samples.

Forensic biology also benefits from using DNA-based identification techniques as disclosed herein. DNA sequences taken from very small biological samples such as tissues, e.g., hair or skin, or body fluids, e.g., blood, saliva, semen, etc., can be amplified using PCR. In one prior art technique, gene sequences amplified from polymorphic loci, such as DQa class II HLA gene, are used in forensic biology to identify individuals. (Erlich, H., PCR Technology, Freeman and Co. (1992).) Once these specific polymorphic loci are amplified, they are digested with one or more restriction enzymes, yielding an identifying set of bands on a Southern blot probed with DNA corresponding to the DQa class II HLA gene. Similarly, polynucleotides of the present invention can be used as polymorphic markers for forensic purposes.

There is also a need for reagents capable of identifying the source of a particular tissue. Such need arises, for example, in forensics when presented with tissue of unknown origin. Appropriate reagents can comprise, for example, DNA probes or primers specific to particular tissue prepared from the sequences of the present invention. Panels of such reagents can identify tissue by species and/or by organ type. In a similar fashion, these reagents can be used to screen tissue cultures for contamination.

In the very least, the polynucleotides of the present invention can be used as molecular weight markers on Southern gels, as diagnostic probes for the presence of a specific mRNA in a particular cell type, as a probe to "subtract-out" known sequences in the process of discovering novel polynucleotides, for selecting and making oligomers for attachment to a "gene chip" or other support, to raise anti-DNA

antibodies using DNA immunization techniques, and as an antigen to elicit an immune response.

Uses of the Polypeptides

5

10

20

25

30

Each of the polypeptides identified herein can be used in numerous ways. The following description should be considered exemplary and utilizes known techniques.

A polypeptide of the present invention can be used to assay protein levels in a biological sample using antibody-based techniques. For example, protein expression in tissues can be studied with classical immunohistological methods. (Jalkanen, M., et al., J. Cell. Biol. 101:976-985 (1985); Jalkanen, M., et al., J. Cell . Biol. 105:3087-3096 (1987).) Other antibody-based methods useful for detecting protein gene expression include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA). Suitable antibody assay labels are known in the art and include enzyme labels, such as, glucose oxidase, and radioisotopes, such as iodine (125I, 121I), carbon (14C), sulfur (35S), tritium (3H), indium (112In), and technetium (99mTc), and fluorescent labels, such as fluorescein and rhodamine, and biotin.

In addition to assaying secreted protein levels in a biological sample, proteins can also be detected in vivo by imaging. Antibody labels or markers for in vivo imaging of protein include those detectable by X-radiography, NMR or ESR. For X-radiography, suitable labels include radioisotopes such as barium or cesium, which emit detectable radiation but are not overtly harmful to the subject. Suitable markers for NMR and ESR include those with a detectable characteristic spin, such as deuterium, which may be incorporated into the antibody by labeling of nutrients for the relevant hybridoma.

A protein-specific antibody or antibody fragment which has been labeled with an appropriate detectable imaging moiety, such as a radioisotope (for example, 131I, 112In, 99mTc), a radio-opaque substance, or a material detectable by nuclear magnetic resonance, is introduced (for example, parenterally, subcutaneously, or intraperitoneally) into the mammal. It will be understood in the art that the size of the subject and the imaging system used will determine the quantity of imaging moiety

WO 99/47540

5

10

15

20

25

30

needed to produce diagnostic images. In the case of a radioisotope moiety, for a human subject, the quantity of radioactivity injected will normally range from about 5 to 20 millicuries of 99mTc. The labeled antibody or antibody fragment will then preferentially accumulate at the location of cells which contain the specific protein. In vivo tumor imaging is described in S.W. Burchiel et al., "Immunopharmacokinetics of Radiolabeled Antibodies and Their Fragments." (Chapter 13 in Tumor Imaging: The Radiochemical Detection of Cancer, S.W. Burchiel and B. A. Rhodes, eds., Masson Publishing Inc. (1982).)

213

PCT/US99/05804

Thus, the invention provides a diagnostic method of a disorder, which involves (a) assaying the expression of a polypeptide of the present invention in cells or body fluid of an individual; (b) comparing the level of gene expression with a standard gene expression level, whereby an increase or decrease in the assayed polypeptide gene expression level compared to the standard expression level is indicative of a disorder.

Moreover, polypeptides of the present invention can be used to treat disease. For example, patients can be administered a polypeptide of the present invention in an effort to replace absent or decreased levels of the polypeptide (e.g., insulin), to supplement absent or decreased levels of a different polypeptide (e.g., hemoglobin S for hemoglobin B), to inhibit the activity of a polypeptide (e.g., an oncogene), to activate the activity of a polypeptide (e.g., by binding to a receptor), to reduce the activity of a membrane bound receptor by competing with it for free ligand (e.g., soluble TNF receptors used in reducing inflammation), or to bring about a desired response (e.g., blood vessel growth).

Similarly, antibodies directed to a polypeptide of the present invention can also be used to treat disease. For example, administration of an antibody directed to a polypeptide of the present invention can bind and reduce overproduction of the polypeptide. Similarly, administration of an antibody can activate the polypeptide, such as by binding to a polypeptide bound to a membrane (receptor).

At the very least, the polypeptides of the present invention can be used as molecular weight markers on SDS-PAGE gels or on molecular sieve gel filtration columns using methods well known to those of skill in the art. Polypeptides can also

be used to raise antibodies, which in turn are used to measure protein expression from a recombinant cell, as a way of assessing transformation of the host cell. Moreover, the polypeptides of the present invention can be used to test the following biological activities.

5

10

20

25

30

Biological Activities

The polynucleotides and polypeptides of the present invention can be used in assays to test for one or more biological activities. If these polynucleotides and polypeptides do exhibit activity in a particular assay, it is likely that these molecules may be involved in the diseases associated with the biological activity. Thus, the polynucleotides and polypeptides could be used to treat the associated disease.

Immune Activity

A polypeptide or polynucleotide of the present invention may be useful in treating deficiencies or disorders of the immune system, by activating or inhibiting the proliferation, differentiation, or mobilization (chemotaxis) of immune cells. Immune cells develop through a process called hematopoiesis, producing myeloid (platelets, red blood cells, neutrophils, and macrophages) and lymphoid (B and T lymphocytes) cells from pluripotent stem cells. The etiology of these immune deficiencies or disorders may be genetic, somatic, such as cancer or some autoimmune disorders, acquired (e.g., by chemotherapy or toxins), or infectious. Moreover, a polynucleotide or polypeptide of the present invention can be used as a marker or detector of a particular immune system disease or disorder.

A polynucleotide or polypeptide of the present invention may be useful in treating or detecting deficiencies or disorders of hematopoietic cells. A polypeptide or polynucleotide of the present invention could be used to increase differentiation and proliferation of hematopoietic cells, including the pluripotent stem cells, in an effort to treat those disorders associated with a decrease in certain (or many) types hematopoietic cells. Examples of immunologic deficiency syndromes include, but are not limited to: blood protein disorders (e.g. agammaglobulinemia, dysgammaglobulinemia), ataxia telangiectasia, common variable immunodeficiency,

WO 99/47540 PCT/US99/05804

215

5

10

15

20

25

30

Digeorge Syndrome, HIV infection, HTLV-BLV infection, leukocyte adhesion deficiency syndrome, lymphopenia, phagocyte bactericidal dysfunction, severe combined immunodeficiency (SCIDs), Wiskott-Aldrich Disorder, anemia, thrombocytopenia, or hemoglobinuria.

Moreover, a polypeptide or polynucleotide of the present invention could also be used to modulate hemostatic (the stopping of bleeding) or thrombolytic activity (clot formation). For example, by increasing hemostatic or thrombolytic activity, a polynucleotide or polypeptide of the present invention could be used to treat blood coagulation disorders (e.g., afibrinogenemia, factor deficiencies), blood platelet disorders (e.g. thrombocytopenia), or wounds resulting from trauma, surgery, or other causes. Alternatively, a polynucleotide or polypeptide of the present invention that can decrease hemostatic or thrombolytic activity could be used to inhibit or dissolve clotting. These molecules could be important in the treatment of heart attacks (infarction), strokes, or scarring.

A polynucleotide or polypeptide of the present invention may also be useful in treating or detecting autoimmune disorders. Many autoimmune disorders result from inappropriate recognition of self as foreign material by immune cells. This inappropriate recognition results in an immune response leading to the destruction of the host tissue. Therefore, the administration of a polypeptide or polynucleotide of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing autoimmune disorders.

Examples of autoimmune disorders that can be treated or detected by the present invention include, but are not limited to: Addison's Disease, hemolytic anemia, antiphospholipid syndrome, rheumatoid arthritis, dermatitis, allergic encephalomyelitis, glomerulonephritis, Goodpasture's Syndrome, Graves' Disease, Multiple Sclerosis, Myasthenia Gravis, Neuritis, Ophthalmia, Bullous Pemphigoid, Pemphigus, Polyendocrinopathies, Purpura, Reiter's Disease, Stiff-Man Syndrome, Autoimmune Thyroiditis, Systemic Lupus Erythematosus, Autoimmune Pulmonary Inflammation, Guillain-Barre Syndrome, insulin dependent diabetes mellitis, and autoimmune inflammatory eye disease.

PCT/US99/05804

Similarly, allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems, may also be treated by a polypeptide or polynucleotide of the present invention. Moreover, these molecules can be used to treat anaphylaxis, hypersensitivity to an antigenic molecule, or blood group incompatibility.

A polynucleotide or polypeptide of the present invention may also be used to treat and/or prevent organ rejection or graft-versus-host disease (GVHD). Organ rejection occurs by host immune cell destruction of the transplanted tissue through an immune response. Similarly, an immune response is also involved in GVHD, but, in this case, the foreign transplanted immune cells destroy the host tissues. The administration of a polypeptide or polynucleotide of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing organ rejection or GVHD.

Similarly, a polypeptide or polynucleotide of the present invention may also be used to modulate inflammation. For example, the polypeptide or polynucleotide may inhibit the proliferation and differentiation of cells involved in an inflammatory response. These molecules can be used to treat inflammatory conditions, both chronic and acute conditions, including inflammation associated with infection (e.g., septic shock, sepsis, or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine induced lung injury, inflammatory bowel disease, Crohn's disease, or resulting from over production of cytokines (e.g., TNF or IL-1.)

25 **Hyperproliferative Disorders**

5

10

15

20

30

A polypeptide or polynucleotide can be used to treat or detect hyperproliferative disorders, including neoplasms. A polypeptide or polynucleotide of the present invention may inhibit the proliferation of the disorder through direct or indirect interactions. Alternatively, a polypeptide or polynucleotide of the present invention may proliferate other cells which can inhibit the hyperproliferative disorder.

10

25

30

For example, by increasing an immune response, particularly increasing antigenic qualities of the hyperproliferative disorder or by proliferating, differentiating, or mobilizing T-cells, hyperproliferative disorders can be treated. This immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, decreasing an immune response may also be a method of treating hyperproliferative disorders, such as a chemotherapeutic agent.

Examples of hyperproliferative disorders that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but are not limited to neoplasms located in the: abdomen, bone, breast, digestive system, liver, pancreas, peritoneum, endocrine glands (adrenal, parathyroid, pituitary, testicles, ovary, thymus, thyroid), eye, head and neck, nervous (central and peripheral), lymphatic system, pelvic, skin, soft tissue, spleen, thoracic, and urogenital.

Similarly, other hyperproliferative disorders can also be treated or detected by

a polynucleotide or polypeptide of the present invention. Examples of such
hyperproliferative disorders include, but are not limited to:
hypergammaglobulinemia, lymphoproliferative disorders, paraproteinemias, purpura,
sarcoidosis, Sezary Syndrome, Waldenstron's Macroglobulinemia, Gaucher's
Disease, histiocytosis, and any other hyperproliferative disease, besides neoplasia,
located in an organ system listed above.

Infectious Disease

A polypeptide or polynucleotide of the present invention can be used to treat or detect infectious agents. For example, by increasing the immune response, particularly increasing the proliferation and differentiation of B and/or T cells, infectious diseases may be treated. The immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, the polypeptide or polynucleotide of the present invention may also directly inhibit the infectious agent, without necessarily eliciting an immune response.

Viruses are one example of an infectious agent that can cause disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the

present invention. Examples of viruses, include, but are not limited to the following DNA and RNA viral families: Arbovirus, Adenoviridae, Arenaviridae, Arterivirus, Birnaviridae, Bunyaviridae, Caliciviridae, Circoviridae, Coronaviridae, Flaviviridae, Hepadnaviridae (Hepatitis), Herpesviridae (such as, Cytomegalovirus, Herpes 5 Simplex, Herpes Zoster), Mononegavirus (e.g., Paramyxoviridae, Morbillivirus, Rhabdoviridae), Orthomyxoviridae (e.g., Influenza), Papovaviridae, Parvoviridae, Picornaviridae, Poxviridae (such as Smallpox or Vaccinia), Reoviridae (e.g., Rotavirus), Retroviridae (HTLV-I, HTLV-II, Lentivirus), and Togaviridae (e.g., Rubivirus). Viruses falling within these families can cause a variety of diseases or 10 symptoms, including, but not limited to: arthritis, bronchiollitis, encephalitis, eye infections (e.g., conjunctivitis, keratitis), chronic fatigue syndrome, hepatitis (A, B, C, E, Chronic Active, Delta), meningitis, opportunistic infections (e.g., AIDS), pneumonia, Burkitt's Lymphoma, chickenpox, hemorrhagic fever, Measles, Mumps, Parainfluenza, Rabies, the common cold, Polio, leukemia, Rubella, sexually 15 transmitted diseases, skin diseases (e.g., Kaposi's, warts), and viremia. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

218

PCT/US99/05804

that can be treated or detected by a polynucleotide or polypeptide of the present 20 invention include, but not limited to, the following Gram-Negative and Gram-positive bacterial families and fungi: Actinomycetales (e.g., Corynebacterium, Mycobacterium, Norcardia), Aspergillosis, Bacillaceae (e.g., Anthrax, Clostridium), Bacteroidaceae, Blastomycosis, Bordetella, Borrelia, Brucellosis, Candidiasis, Campylobacter, Coccidioidomycosis, Cryptococcosis, Dermatocycoses, 25 Enterobacteriaceae (Klebsiella, Salmonella, Serratia, Yersinia), Erysipelothrix, Helicobacter, Legionellosis, Leptospirosis, Listeria, Mycoplasmatales, Neisseriaceae (e.g., Acinetobacter, Gonorrhea, Menigococcal), Pasteurellacea Infections (e.g., Actinobacillus, Heamophilus, Pasteurella), Pseudomonas, Rickettsiaceae, Chlamydiaceae, Syphilis, and Staphylococcal. These bacterial or fungal families can 30 cause the following diseases or symptoms, including, but not limited to: bacteremia, endocarditis, eye infections (conjunctivitis, tuberculosis, uveitis), gingivitis,

Similarly, bacterial or fungal agents that can cause disease or symptoms and

opportunistic infections (e.g., AIDS related infections), paronychia, prosthesis-related infections, Reiter's Disease, respiratory tract infections, such as Whooping Cough or Empyema, sepsis, Lyme Disease, Cat-Scratch Disease, Dysentery, Paratyphoid Fever, food poisoning, Typhoid, pneumonia, Gonorrhea, meningitis, Chlamydia, Syphilis, Diphtheria, Leprosy, Paratuberculosis, Tuberculosis, Lupus, Botulism, gangrene, tetanus, impetigo, Rheumatic Fever, Scarlet Fever, sexually transmitted diseases, skin diseases (e.g., cellulitis, dermatocycoses), toxemia, urinary tract infections, wound infections. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Moreover, parasitic agents causing disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following families: Amebiasis, Babesiosis, Coccidiosis, Cryptosporidiosis, Dientamoebiasis, Dourine, Ectoparasitic, Giardiasis, Helminthiasis, Leishmaniasis, Theileriasis, Toxoplasmosis, Trypanosomiasis, and Trichomonas. These parasites can cause a variety of diseases or symptoms, including, but not limited to: Scabies, Trombiculiasis, eye infections, intestinal disease (e.g., dysentery, giardiasis), liver disease, lung disease, opportunistic infections (e.g., AIDS related), Malaria, pregnancy complications, and toxoplasmosis. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Preferably, treatment using a polypeptide or polynucleotide of the present invention could either be by administering an effective amount of a polypeptide to the patient, or by removing cells from the patient, supplying the cells with a polynucleotide of the present invention, and returning the engineered cells to the patient (ex vivo therapy). Moreover, the polypeptide or polynucleotide of the present invention can be used as an antigen in a vaccine to raise an immune response against infectious disease.

Regeneration

25

30

5

A polynucleotide or polypeptide of the present invention can be used to differentiate, proliferate, and attract cells, leading to the regeneration of tissues. (See,

WO 99/47540 PCT/US99/05804

220

Science 276:59-87 (1997).) The regeneration of tissues could be used to repair, replace, or protect tissue damaged by congenital defects, trauma (wounds, burns, incisions, or ulcers), age, disease (e.g. osteoporosis, osteocarthritis, periodontal disease, liver failure), surgery, including cosmetic plastic surgery, fibrosis, reperfusion injury, or systemic cytokine damage.

Tissues that could be regenerated using the present invention include organs (e.g., pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac), vasculature (including vascular and lymphatics), nervous, hematopoietic, and skeletal (bone, cartilage, tendon, and ligament) tissue. Preferably, regeneration occurs without or decreased scarring. Regeneration also may include angiogenesis.

Moreover, a polynucleotide or polypeptide of the present invention may increase regeneration of tissues difficult to heal. For example, increased tendon/ligament regeneration would quicken recovery time after damage. A polynucleotide or polypeptide of the present invention could also be used prophylactically in an effort to avoid damage. Specific diseases that could be treated include of tendinitis, carpal tunnel syndrome, and other tendon or ligament defects. A further example of tissue regeneration of non-healing wounds includes pressure ulcers, ulcers associated with vascular insufficiency, surgical, and traumatic wounds.

Similarly, nerve and brain tissue could also be regenerated by using a polynucleotide or polypeptide of the present invention to proliferate and differentiate nerve cells. Diseases that could be treated using this method include central and peripheral nervous system diseases, neuropathies, or mechanical and traumatic disorders (e.g., spinal cord disorders, head trauma, cerebrovascular disease, and stoke). Specifically, diseases associated with peripheral nerve injuries, peripheral neuropathy (e.g., resulting from chemotherapy or other medical therapies), localized neuropathies, and central nervous system diseases (e.g., Alzheimer's disease, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome), could all be treated using the polynucleotide or polypeptide of the present invention.

30

5

10

15

20

25

Chemotaxis

10

15

20

25

30

A polynucleotide or polypeptide of the present invention may have chemotaxis activity. A chemotaxic molecule attracts or mobilizes cells (e.g., monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells) to a particular site in the body, such as inflammation, infection, or site of hyperproliferation. The mobilized cells can then fight off and/or heal the particular trauma or abnormality.

A polynucleotide or polypeptide of the present invention may increase chemotaxic activity of particular cells. These chemotactic molecules can then be used to treat inflammation, infection, hyperproliferative disorders, or any immune system disorder by increasing the number of cells targeted to a particular location in the body. For example, chemotaxic molecules can be used to treat wounds and other trauma to tissues by attracting immune cells to the injured location. Chemotactic molecules of the present invention can also attract fibroblasts, which can be used to treat wounds.

It is also contemplated that a polynucleotide or polypeptide of the present invention may inhibit chemotactic activity. These molecules could also be used to treat disorders. Thus, a polynucleotide or polypeptide of the present invention could be used as an inhibitor of chemotaxis.

Binding Activity

A polypeptide of the present invention may be used to screen for molecules that bind to the polypeptide or for molecules to which the polypeptide binds. The binding of the polypeptide and the molecule may activate (agonist), increase, inhibit (antagonist), or decrease activity of the polypeptide or the molecule bound. Examples of such molecules include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

Preferably, the molecule is closely related to the natural ligand of the polypeptide, e.g., a fragment of the ligand, or a natural substrate, a ligand, a structural or functional mimetic. (See, Coligan et al., Current Protocols in Immunology 1(2):Chapter 5 (1991).) Similarly, the molecule can be closely related to the natural receptor to which the polypeptide binds, or at least, a fragment of the receptor capable

10

15

20

25

30

of being bound by the polypeptide (e.g., active site). In either case, the molecule can be rationally designed using known techniques.

Preferably, the screening for these molecules involves producing appropriate cells which express the polypeptide, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or *E. coli*. Cells expressing the polypeptide (or cell membrane containing the expressed polypeptide) are then preferably contacted with a test compound potentially containing the molecule to observe binding, stimulation, or inhibition of activity of either the polypeptide or the molecule.

The assay may simply test binding of a candidate compound to the polypeptide, wherein binding is detected by a label, or in an assay involving competition with a labeled competitor. Further, the assay may test whether the candidate compound results in a signal generated by binding to the polypeptide.

Alternatively, the assay can be carried out using cell-free preparations, polypeptide/molecule affixed to a solid support, chemical libraries, or natural product mixtures. The assay may also simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide, measuring polypeptide/molecule activity or binding, and comparing the polypeptide/molecule activity or binding to a standard.

Preferably, an ELISA assay can measure polypeptide level or activity in a sample (e.g., biological sample) using a monoclonal or polyclonal antibody. The antibody can measure polypeptide level or activity by either binding, directly or indirectly, to the polypeptide or by competing with the polypeptide for a substrate.

All of these above assays can be used as diagnostic or prognostic markers. The molecules discovered using these assays can be used to treat disease or to bring about a particular result in a patient (e.g., blood vessel growth) by activating or inhibiting the polypeptide/molecule. Moreover, the assays can discover agents which may inhibit or enhance the production of the polypeptide from suitably manipulated cells or tissues.

Therefore, the invention includes a method of identifying compounds which bind to a polypeptide of the invention comprising the steps of: (a) incubating a

candidate binding compound with a polypeptide of the invention; and (b) determining if binding has occurred. Moreover, the invention includes a method of identifying agonists/antagonists comprising the steps of: (a) incubating a candidate compound with a polypeptide of the invention, (b) assaying a biological activity, and (b) determining if a biological activity of the polypeptide has been altered.

Other Activities

5

10

15

20

25

30

A polypeptide or polynucleotide of the present invention may also increase or decrease the differentiation or proliferation of embryonic stem cells, besides, as discussed above, hematopoietic lineage.

A polypeptide or polynucleotide of the present invention may also be used to modulate mammalian characteristics, such as body height, weight, hair color, eye color, skin, percentage of adipose tissue, pigmentation, size, and shape (e.g., cosmetic surgery). Similarly, a polypeptide or polynucleotide of the present invention may be used to modulate mammalian metabolism affecting catabolism, anabolism, processing, utilization, and storage of energy.

A polypeptide or polynucleotide of the present invention may be used to change a mammal's mental state or physical state by influencing biorhythms, caricadic rhythms, depression (including depressive disorders), tendency for violence, tolerance for pain, reproductive capabilities (preferably by Activin or Inhibin-like activity), hormonal or endocrine levels, appetite, libido, memory, stress, or other cognitive qualities.

A polypeptide or polynucleotide of the present invention may also be used as a food additive or preservative, such as to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional components.

Other Preferred Embodiments

Other preferred embodiments of the claimed invention include an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95%

5

10

15

20

25

30

224

PCT/US99/05804

identical to a sequence of at least about 50 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Clone Sequence and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Start Codon and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Similarly preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 150 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

Further preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 500 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

A further preferred embodiment is a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the nucleotide sequence of SEQ ID NO:X beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

10

15

20

25

30

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence of SEQ ID NO:X.

Also preferred is an isolated nucleic acid molecule which hybridizes under stringent hybridization conditions to a nucleic acid molecule, wherein said nucleic acid molecule which hybridizes does not hybridize under stringent hybridization conditions to a nucleic acid molecule having a nucleotide sequence consisting of only A residues or of only T residues.

Also preferred is a composition of matter comprising a DNA molecule which comprises a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the material deposited with the American Type Culture Collection and given the ATCC Deposit Number shown in Table 1 for said cDNA Clone Identifier.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in the nucleotide sequence of a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the deposit given the ATCC Deposit Number shown in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said sequence of at least 50 contiguous nucleotides is included in the nucleotide sequence of the complete open reading frame sequence encoded by said human cDNA clone.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 150 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 500 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence encoded by said human cDNA clone.

WO 99/47540 PCT/US99/05804

A further preferred embodiment is a method for detecting in a biological sample a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing a nucleotide sequence of at least one nucleic acid molecule in said sample with a sequence selected from said group and determining whether the sequence of said nucleic acid molecule in said sample is at least 95% identical to said selected sequence.

10

15

20

25

30

226

Also preferred is the above method wherein said step of comparing sequences comprises determining the extent of nucleic acid hybridization between nucleic acid molecules in said sample and a nucleic acid molecule comprising said sequence selected from said group. Similarly, also preferred is the above method wherein said step of comparing sequences is performed by comparing the nucleotide sequence determined from a nucleic acid molecule in said sample with said sequence selected from said group. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

A further preferred embodiment is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting nucleic acid molecules in said sample, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

The method for identifying the species, tissue or cell type of a biological sample can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least

10

15

20

25

30

one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject nucleic acid molecules, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

The method for diagnosing a pathological condition can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a composition of matter comprising isolated nucleic acid molecules wherein the nucleotide sequences of said nucleic acid molecules comprise a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1.

10

15

20

25

30

Also preferred is a polypeptide, wherein said sequence of contiguous amino acids is included in the amino acid sequence of SEQ ID NO:Y in the range of positions beginning with the residue at about the position of the First Amino Acid of the Secreted Portion and ending with the residue at about the Last Amino Acid of the Open Reading Frame as set forth for SEQ ID NO:Y in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the complete amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is a polypeptide wherein said sequence of contiguous amino acids is included in the amino acid sequence of a secreted portion of the secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in

the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

5

10

15

20

25

30

Further preferred is an isolated antibody which binds specifically to a polypeptide comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is a method for detecting in a biological sample a polypeptide comprising an amino acid sequence which is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group and determining whether the sequence of at least 10 contiguous amino acids.

Also preferred is the above method wherein said step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group comprises determining the extent of specific binding of polypeptides in said sample to an antibody which binds specifically to a polypeptide

10

15

20

25

30

PCT/US99/05804

comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is the above method wherein said step of comparing sequences is performed by comparing the amino acid sequence determined from a polypeptide molecule in said sample with said sequence selected from said group.

Also preferred is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting polypeptide molecules in said sample, if any, comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is the above method for identifying the species, tissue or cell type of a biological sample, which method comprises a step of detecting polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the above group.

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino

PCT/US99/05804

acid sequence of SEQ ID NO: Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

In any of these methods, the step of detecting said polypeptide molecules includes using an antibody.

5

10

15

20

25

30

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a nucleotide sequence encoding a polypeptide wherein said polypeptide comprises an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said nucleotide sequence encoding a polypeptide has been optimized for expression of said polypeptide in a prokaryotic host.

Also preferred is an isolated nucleic acid molecule, wherein said polypeptide comprises an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is a method of making a recombinant vector comprising inserting any of the above isolated nucleic acid molecule into a vector. Also preferred is the recombinant vector produced by this method. Also preferred is a method of making a recombinant host cell comprising introducing the vector into a host cell, as well as the recombinant host cell produced by this method.

Also preferred is a method of making an isolated polypeptide comprising culturing this recombinant host cell under conditions such that said polypeptide is

expressed and recovering said polypeptide. Also preferred is this method of making an isolated polypeptide, wherein said recombinant host cell is a eukaryotic cell and said polypeptide is a secreted portion of a human secreted protein comprising an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y beginning with the residue at the position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y wherein Y is an integer set forth in Table 1 and said position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y is defined in Table 1; and an amino acid sequence of a secreted portion of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The isolated polypeptide produced by this method is also preferred.

Also preferred is a method of treatment of an individual in need of an increased level of a secreted protein activity, which method comprises administering to such an individual a pharmaceutical composition comprising an amount of an isolated polypeptide, polynucleotide, or antibody of the claimed invention effective to increase the level of said protein activity in said individual.

Having generally described the invention, the same will be more readily understood by reference to the following examples, which are provided by way of illustration and are not intended as limiting.

20

25

30

5

10

15

Examples

Example 1: Isolation of a Selected cDNA Clone From the Deposited Sample

Each cDNA clone in a cited ATCC deposit is contained in a plasmid vector. Table 1 identifies the vectors used to construct the cDNA library from which each clone was isolated. In many cases, the vector used to construct the library is a phage vector from which a plasmid has been excised. The table immediately below correlates the related plasmid for each phage vector used in constructing the cDNA library. For example, where a particular clone is identified in Table 1 as being isolated in the vector "Lambda Zap," the corresponding deposited clone is in "pBluescript."

	Vector Used to Construct Library	Corresponding Deposited
	<u>Plasmid</u>	
	Lambda Zap	pBluescript (pBS)
	Uni-Zap XR	pBluescript (pBS)
5	Zap Express	pBK
	lafmid BA	plafmid BA
	pSport1	pSport1
	pCMVSport 2.0	pCMVSport 2.0
	pCMVSport 3.0	pCMVSport 3.0
10	pCR [®] 2.1	pCR®2.1

Vectors Lambda Zap (U.S. Patent Nos. 5,128,256 and 5,286,636), Uni-Zap XR (U.S. Patent Nos. 5,128, 256 and 5,286,636), Zap Express (U.S. Patent Nos. 5,128,256 and 5,286,636), pBluescript (pBS) (Short, J. M. et al., Nucleic Acids Res. 16:7583-7600 (1988); Alting-Mees, M. A. and Short, J. M., Nucleic Acids Res. 15 17:9494 (1989)) and pBK (Alting-Mees, M. A. et al., Strategies 5:58-61 (1992)) are commercially available from Stratagene Cloning Systems, Inc., 11011 N. Torrey Pines Road, La Jolla, CA, 92037. pBS contains an ampicillin resistance gene and pBK contains a neomycin resistance gene. Both can be transformed into E. coli strain XL-1 Blue, also available from Stratagene. pBS comes in 4 forms SK+, SK-, KS+ and KS. The S and K refers to the orientation of the polylinker to the T7 and T3 20 primer sequences which flank the polylinker region ("S" is for SacI and "K" is for KpnI which are the first sites on each respective end of the linker). "+" or "-" refer to the orientation of the f1 origin of replication ("ori"), such that in one orientation, single stranded rescue initiated from the f1 ori generates sense strand DNA and in the 25 other, antisense.

Vectors pSport1, pCMVSport 2.0 and pCMVSport 3.0, were obtained from Life Technologies, Inc., P. O. Box 6009, Gaithersburg, MD 20897. All Sport vectors contain an ampicillin resistance gene and may be transformed into E. coli strain DH10B, also available from Life Technologies. (See, for instance, Gruber, C. E., et al., Focus 15:59 (1993).) Vector lafmid BA (Bento Soares, Columbia University, NY) contains an ampicillin resistance gene and can be transformed into E. coli strain

WO 99/47540 PCT/US99/05804

234

XL-1 Blue. Vector pCR[®]2.1, which is available from Invitrogen, 1600 Faraday Avenue, Carlsbad, CA 92008, contains an ampicillin resistance gene and may be transformed into E. coli strain DH10B, available from Life Technologies. (See, for instance, Clark, J. M., Nuc. Acids Res. 16:9677-9686 (1988) and Mead, D. et al., Bio/Technology 9: (1991).) Preferably, a polynucleotide of the present invention does not comprise the phage vector sequences identified for the particular clone in Table 1, as well as the corresponding plasmid vector sequences designated above.

The deposited material in the sample assigned the ATCC Deposit Number cited in Table 1 for any given cDNA clone also may contain one or more additional plasmids, each comprising a cDNA clone different from that given clone. Thus, deposits sharing the same ATCC Deposit Number contain at least a plasmid for each cDNA clone identified in Table 1. Typically, each ATCC deposit sample cited in Table 1 comprises a mixture of approximately equal amounts (by weight) of about 50 plasmid DNAs, each containing a different cDNA clone; but such a deposit sample may include plasmids for more or less than 50 cDNA clones, up to about 500 cDNA clones.

10

15

20

25

30

Two approaches can be used to isolate a particular clone from the deposited sample of plasmid DNAs cited for that clone in Table 1. First, a plasmid is directly isolated by screening the clones using a polynucleotide probe corresponding to SEQ ID NO:X.

Particularly, a specific polynucleotide with 30-40 nucleotides is synthesized using an Applied Biosystems DNA synthesizer according to the sequence reported. The oligonucleotide is labeled, for instance, with ³²P-γ-ATP using T4 polynucleotide kinase and purified according to routine methods. (E.g., Maniatis et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring, NY (1982).) The plasmid mixture is transformed into a suitable host, as indicated above (such as XL-1 Blue (Stratagene)) using techniques known to those of skill in the art, such as those provided by the vector supplier or in related publications or patents cited above. The transformants are plated on 1.5% agar plates (containing the appropriate selection agent, e.g., ampicillin) to a density of about 150 transformants (colonies) per plate. These plates are screened using Nylon membranes according to routine methods for

5

10

15

20

25

30

bacterial colony screening (e.g., Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd Edit., (1989), Cold Spring Harbor Laboratory Press, pages 1.93 to 1.104), or other techniques known to those of skill in the art.

Alternatively, two primers of 17-20 nucleotides derived from both ends of the SEQ ID NO:X (i.e., within the region of SEQ ID NO:X bounded by the 5' NT and the 3' NT of the clone defined in Table 1) are synthesized and used to amplify the desired cDNA using the deposited cDNA plasmid as a template. The polymerase chain reaction is carried out under routine conditions, for instance, in 25 µl of reaction mixture with 0.5 ug of the above cDNA template. A convenient reaction mixture is 1.5-5 mM MgCl₂, 0.01% (w/v) gelatin, 20 µM each of dATP, dCTP, dGTP, dTTP, 25 pmol of each primer and 0.25 Unit of Taq polymerase. Thirty five cycles of PCR (denaturation at 94°C for 1 min; annealing at 55°C for 1 min; elongation at 72°C for 1 min) are performed with a Perkin-Elmer Cetus automated thermal cycler. The amplified product is analyzed by agarose gel electrophoresis and the DNA band with expected molecular weight is excised and purified. The PCR product is verified to be the selected sequence by subcloning and sequencing the DNA product.

Several methods are available for the identification of the 5' or 3' non-coding portions of a gene which may not be present in the deposited clone. These methods include but are not limited to, filter probing, clone enrichment using specific probes, and protocols similar or identical to 5' and 3' "RACE" protocols which are well known in the art. For instance, a method similar to 5' RACE is available for generating the missing 5' end of a desired full-length transcript. (Fromont-Racine et al., Nucleic Acids Res. 21(7):1683-1684 (1993).)

Briefly, a specific RNA oligonucleotide is ligated to the 5' ends of a population of RNA presumably containing full-length gene RNA transcripts. A primer set containing a primer specific to the ligated RNA oligonucleotide and a primer specific to a known sequence of the gene of interest is used to PCR amplify the 5' portion of the desired full-length gene. This amplified product may then be sequenced and used to generate the full length gene.

This above method starts with total RNA isolated from the desired source, although poly-A+ RNA can be used. The RNA preparation can then be treated with

5

10

15

20

25

30

phosphatase if necessary to eliminate 5' phosphate groups on degraded or damaged RNA which may interfere with the later RNA ligase step. The phosphatase should then be inactivated and the RNA treated with tobacco acid pyrophosphatase in order to remove the cap structure present at the 5' ends of messenger RNAs. This reaction leaves a 5' phosphate group at the 5' end of the cap cleaved RNA which can then be ligated to an RNA oligonucleotide using T4 RNA ligase.

This modified RNA preparation is used as a template for first strand cDNA synthesis using a gene specific oligonucleotide. The first strand synthesis reaction is used as a template for PCR amplification of the desired 5' end using a primer specific to the ligated RNA oligonucleotide and a primer specific to the known sequence of the gene of interest. The resultant product is then sequenced and analyzed to confirm that the 5' end sequence belongs to the desired gene.

Example 2: Isolation of Genomic Clones Corresponding to a Polynucleotide

A human genomic P1 library (Genomic Systems, Inc.) is screened by PCR using primers selected for the cDNA sequence corresponding to SEQ ID NO:X., according to the method described in Example 1. (See also, Sambrook.)

Example 3: Tissue Distribution of Polypeptide

Tissue distribution of mRNA expression of polynucleotides of the present invention is determined using protocols for Northern blot analysis, described by, among others, Sambrook et al. For example, a cDNA probe produced by the method described in Example 1 is labeled with P³² using the rediprimeTM DNA labeling system (Amersham Life Science), according to manufacturer's instructions. After labeling, the probe is purified using CHROMA SPIN-100TM column (Clontech Laboratories, Inc.), according to manufacturer's protocol number PT1200-1. The purified labeled probe is then used to examine various human tissues for mRNA expression.

Multiple Tissue Northern (MTN) blots containing various human tissues (H) or human immune system tissues (IM) (Clontech) are examined with the labeled probe using ExpressHybTM hybridization solution (Clontech) according to

10

15

20

25

manufacturer's protocol number PT1190-1. Following hybridization and washing, the blots are mounted and exposed to film at -70°C overnight, and the films developed according to standard procedures.

5 Example 4: Chromosomal Mapping of the Polynucleotides

An oligonucleotide primer set is designed according to the sequence at the 5' end of SEQ ID NO:X. This primer preferably spans about 100 nucleotides. This primer set is then used in a polymerase chain reaction under the following set of conditions: 30 seconds, 95°C; 1 minute, 56°C; 1 minute, 70°C. This cycle is repeated 32 times followed by one 5 minute cycle at 70°C. Human, mouse, and hamster DNA is used as template in addition to a somatic cell hybrid panel containing individual chromosomes or chromosome fragments (Bios, Inc). The reactions is analyzed on either 8% polyacrylamide gels or 3.5 % agarose gels. Chromosome mapping is determined by the presence of an approximately 100 bp PCR fragment in the particular somatic cell hybrid.

Example 5: Bacterial Expression of a Polypeptide

A polynucleotide encoding a polypeptide of the present invention is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' ends of the DNA sequence, as outlined in Example 1, to synthesize insertion fragments. The primers used to amplify the cDNA insert should preferably contain restriction sites, such as BamHI and XbaI, at the 5' end of the primers in order to clone the amplified product into the expression vector. For example, BamHI and XbaI correspond to the restriction enzyme sites on the bacterial expression vector pQE-9. (Qiagen, Inc.,

Chatsworth, CA). This plasmid vector encodes antibiotic resistance (Amp^r), a bacterial origin of replication (ori), an IPTG-regulatable promoter/operator (P/O), a ribosome binding site (RBS), a 6-histidine tag (6-His), and restriction enzyme cloning sites.

The pQE-9 vector is digested with BarnHI and XbaI and the amplified

fragment is ligated into the pQE-9 vector maintaining the reading frame initiated at the bacterial RBS. The ligation mixture is then used to transform the E. coli strain

10

15

20

25

30

M15/rep4 (Qiagen, Inc.) which contains multiple copies of the plasmid pREP4, which expresses the lacI repressor and also confers kanamycin resistance (Kan^r). Transformants are identified by their ability to grow on LB plates and ampicillin/kanamycin resistant colonies are selected. Plasmid DNA is isolated and confirmed by restriction analysis.

Clones containing the desired constructs are grown overnight (O/N) in liquid culture in LB media supplemented with both Amp (100 ug/ml) and Kan (25 ug/ml). The O/N culture is used to inoculate a large culture at a ratio of 1:100 to 1:250. The cells are grown to an optical density 600 (O.D.⁶⁰⁰) of between 0.4 and 0.6. IPTG (Isopropyl-B-D-thiogalacto pyranoside) is then added to a final concentration of 1 mM. IPTG induces by inactivating the lacI repressor, clearing the P/O leading to increased gene expression.

Cells are grown for an extra 3 to 4 hours. Cells are then harvested by centrifugation (20 mins at 6000Xg). The cell pellet is solubilized in the chaotropic agent 6 Molar Guanidine HCl by stirring for 3-4 hours at 4°C. The cell debris is removed by centrifugation, and the supernatant containing the polypeptide is loaded onto a nickel-nitrilo-tri-acetic acid ("Ni-NTA") affinity resin column (available from QIAGEN, Inc., *supra*). Proteins with a 6 x His tag bind to the Ni-NTA resin with high affinity and can be purified in a simple one-step procedure (for details see: The QIAexpressionist (1995) QIAGEN, Inc., *supra*).

Briefly, the supernatant is loaded onto the column in 6 M guanidine-HCl, pH 8, the column is first washed with 10 volumes of 6 M guanidine-HCl, pH 8, then washed with 10 volumes of 6 M guanidine-HCl pH 6, and finally the polypeptide is eluted with 6 M guanidine-HCl, pH 5.

The purified protein is then renatured by dialyzing it against phosphate-buffered saline (PBS) or 50 mM Na-acetate, pH 6 buffer plus 200 mM NaCl.

Alternatively, the protein can be successfully refolded while immobilized on the Ni-NTA column. The recommended conditions are as follows: renature using a linear 6M-1M urea gradient in 500 mM NaCl, 20% glycerol, 20 mM Tris/HCl pH 7.4, containing protease inhibitors. The renaturation should be performed over a period of 1.5 hours or more. After renaturation the proteins are eluted by the addition of 250

5

10

15

20

25

30

PCT/US99/05804

mM immidazole. Immidazole is removed by a final dialyzing step against PBS or 50 mM sodium acetate pH 6 buffer plus 200 mM NaCl. The purified protein is stored at 4°C or frozen at -80°C.

In addition to the above expression vector, the present invention further includes an expression vector comprising phage operator and promoter elements operatively linked to a polynucleotide of the present invention, called pHE4a. (ATCC Accession Number 209645, deposited on February 25, 1998.) This vector contains:

1) a neomycinphosphotransferase gene as a selection marker, 2) an E. coli origin of replication, 3) a T5 phage promoter sequence, 4) two lac operator sequences, 5) a Shine-Delgarno sequence, and 6) the lactose operon repressor gene (lacIq). The origin of replication (oriC) is derived from pUC19 (LTI, Gaithersburg, MD). The promoter sequence and operator sequences are made synthetically.

DNA can be inserted into the pHEa by restricting the vector with NdeI and XbaI, BamHI, XhoI, or Asp718, running the restricted product on a gel, and isolating the larger fragment (the stuffer fragment should be about 310 base pairs). The DNA insert is generated according to the PCR protocol described in Example 1, using PCR primers having restriction sites for NdeI (5' primer) and XbaI, BamHI, XhoI, or Asp718 (3' primer). The PCR insert is gel purified and restricted with compatible enzymes. The insert and vector are ligated according to standard protocols.

The engineered vector could easily be substituted in the above protocol to express protein in a bacterial system.

Example 6: Purification of a Polypeptide from an Inclusion Body

The following alternative method can be used to purify a polypeptide expressed in *E coli* when it is present in the form of inclusion bodies. Unless otherwise specified, all of the following steps are conducted at 4-10°C.

Upon completion of the production phase of the *E. coli* fermentation, the cell culture is cooled to 4-10°C and the cells harvested by continuous centrifugation at 15,000 rpm (Heraeus Sepatech). On the basis of the expected yield of protein per unit weight of cell paste and the amount of purified protein required, an appropriate amount of cell paste, by weight, is suspended in a buffer solution containing 100 mM

10

15

20

25

30

PCT/US99/05804

240

Tris, 50 mM EDTA, pH 7.4. The cells are dispersed to a homogeneous suspension using a high shear mixer.

The cells are then lysed by passing the solution through a microfluidizer (Microfuidics, Corp. or APV Gaulin, Inc.) twice at 4000-6000 psi. The homogenate is then mixed with NaCl solution to a final concentration of 0.5 M NaCl, followed by centrifugation at 7000 xg for 15 min. The resultant pellet is washed again using 0.5M NaCl, 100 mM Tris, 50 mM EDTA, pH 7.4.

The resulting washed inclusion bodies are solubilized with 1.5 M guanidine hydrochloride (GuHCl) for 2-4 hours. After 7000 xg centrifugation for 15 min., the pellet is discarded and the polypeptide containing supernatant is incubated at 4°C overnight to allow further GuHCl extraction.

Following high speed centrifugation (30,000 xg) to remove insoluble particles, the GuHCl solubilized protein is refolded by quickly mixing the GuHCl extract with 20 volumes of buffer containing 50 mM sodium, pH 4.5, 150 mM NaCl, 2 mM EDTA by vigorous stirring. The refolded diluted protein solution is kept at 4°C without mixing for 12 hours prior to further purification steps.

To clarify the refolded polypeptide solution, a previously prepared tangential filtration unit equipped with 0.16 µm membrane filter with appropriate surface area (e.g., Filtron), equilibrated with 40 mM sodium acetate, pH 6.0 is employed. The filtered sample is loaded onto a cation exchange resin (e.g., Poros HS-50, Perseptive Biosystems). The column is washed with 40 mM sodium acetate, pH 6.0 and eluted with 250 mM, 500 mM, 1000 mM, and 1500 mM NaCl in the same buffer, in a stepwise manner. The absorbance at 280 nm of the effluent is continuously monitored. Fractions are collected and further analyzed by SDS-PAGE.

Fractions containing the polypeptide are then pooled and mixed with 4 volumes of water. The diluted sample is then loaded onto a previously prepared set of tandem columns of strong anion (Poros HQ-50, Perseptive Biosystems) and weak anion (Poros CM-20, Perseptive Biosystems) exchange resins. The columns are equilibrated with 40 mM sodium acetate, pH 6.0. Both columns are washed with 40 mM sodium acetate, pH 6.0, 200 mM NaCl. The CM-20 column is then eluted using a 10 column volume linear gradient ranging from 0.2 M NaCl, 50 mM sodium

acetate, pH 6.0 to 1.0 M NaCl, 50 mM sodium acetate, pH 6.5. Fractions are collected under constant A₂₈₀ monitoring of the effluent. Fractions containing the polypeptide (determined, for instance, by 16% SDS-PAGE) are then pooled.

The resultant polypeptide should exhibit greater than 95% purity after the above refolding and purification steps. No major contaminant bands should be observed from Commassie blue stained 16% SDS-PAGE gel when 5 µg of purified protein is loaded. The purified protein can also be tested for endotoxin/LPS contamination, and typically the LPS content is less than 0.1 ng/ml according to LAL assays.

10

15

20

25

30

5

Example 7: Cloning and Expression of a Polypeptide in a Baculovirus Expression System

In this example, the plasmid shuttle vector pA2 is used to insert a polynucleotide into a baculovirus to express a polypeptide. This expression vector contains the strong polyhedrin promoter of the *Autographa californica* nuclear polyhedrosis virus (AcMNPV) followed by convenient restriction sites such as BamHI, Xba I and Asp718. The polyadenylation site of the simian virus 40 ("SV40") is used for efficient polyadenylation. For easy selection of recombinant virus, the plasmid contains the beta-galactosidase gene from *E. coli* under control of a weak Drosophila promoter in the same orientation, followed by the polyadenylation signal of the polyhedrin gene. The inserted genes are flanked on both sides by viral sequences for cell-mediated homologous recombination with wild-type viral DNA to generate a viable virus that express the cloned polynucleotide.

Many other baculovirus vectors can be used in place of the vector above, such as pAc373, pVL941, and pAcIM1, as one skilled in the art would readily appreciate, as long as the construct provides appropriately located signals for transcription, translation, secretion and the like, including a signal peptide and an in-frame AUG as required. Such vectors are described, for instance, in Luckow et al., Virology 170:31-39 (1989).

Specifically, the cDNA sequence contained in the deposited clone, including the AUG initiation codon and the naturally associated leader sequence identified in

Table 1, is amplified using the PCR protocol described in Example 1. If the naturally occurring signal sequence is used to produce the secreted protein, the pA2 vector does not need a second signal peptide. Alternatively, the vector can be modified (pA2 GP) to include a baculovirus leader sequence, using the standard methods described in Summers et al., "A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures," Texas Agricultural Experimental Station Bulletin No. 1555 (1987).

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

10

15

20

The plasmid is digested with the corresponding restriction enzymes and optionally, can be dephosphorylated using calf intestinal phosphatase, using routine procedures known in the art. The DNA is then isolated from a 1% agarose gel using a commercially available kit ("Geneclean" BIO 101 Inc., La Jolla, Ca.).

The fragment and the dephosphorylated plasmid are ligated together with T4 DNA ligase. *E. coli* HB101 or other suitable *E. coli* hosts such as XL-1 Blue (Stratagene Cloning Systems, La Jolla, CA) cells are transformed with the ligation mixture and spread on culture plates. Bacteria containing the plasmid are identified by digesting DNA from individual colonies and analyzing the digestion product by gel electrophoresis. The sequence of the cloned fragment is confirmed by DNA sequencing.

Five μg of a plasmid containing the polynucleotide is co-transfected with 1.0 μg of a commercially available linearized baculovirus DNA ("BaculoGoldTM

25 baculovirus DNA", Pharmingen, San Diego, CA), using the lipofection method described by Felgner et al., Proc. Natl. Acad. Sci. USA 84:7413-7417 (1987). One μg of BaculoGoldTM virus DNA and 5 μg of the plasmid are mixed in a sterile well of a microtiter plate containing 50 μl of serum-free Grace's medium (Life Technologies Inc., Gaithersburg, MD). Afterwards, 10 μl Lipofectin plus 90 μl Grace's medium are added, mixed and incubated for 15 minutes at room temperature. Then the transfection mixture is added drop-wise to Sf9 insect cells (ATCC CRL 1711) seeded

10

15

20

25

30

in a 35 mm tissue culture plate with 1 ml Grace's medium without serum. The plate is then incubated for 5 hours at 27° C. The transfection solution is then removed from the plate and 1 ml of Grace's insect medium supplemented with 10% fetal calf serum is added. Cultivation is then continued at 27° C for four days.

243

After four days the supernatant is collected and a plaque assay is performed, as described by Summers and Smith, *supra*. An agarose gel with "Blue Gal" (Life Technologies Inc., Gaithersburg) is used to allow easy identification and isolation of gal-expressing clones, which produce blue-stained plaques. (A detailed description of a "plaque assay" of this type can also be found in the user's guide for insect cell culture and baculovirology distributed by Life Technologies Inc., Gaithersburg, page 9-10.) After appropriate incubation, blue stained plaques are picked with the tip of a micropipettor (e.g., Eppendorf). The agar containing the recombinant viruses is then resuspended in a microcentrifuge tube containing 200 µl of Grace's medium and the suspension containing the recombinant baculovirus is used to infect Sf9 cells seeded in 35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then they are stored at 4° C.

To verify the expression of the polypeptide, Sf9 cells are grown in Grace's medium supplemented with 10% heat-inactivated FBS. The cells are infected with the recombinant baculovirus containing the polynucleotide at a multiplicity of infection ("MOI") of about 2. If radiolabeled proteins are desired, 6 hours later the medium is removed and is replaced with SF900 II medium minus methionine and cysteine (available from Life Technologies Inc., Rockville, MD). After 42 hours, 5 μCi of ³⁵S-methionine and 5 μCi ³⁵S-cysteine (available from Amersham) are added. The cells are further incubated for 16 hours and then are harvested by centrifugation. The proteins in the supernatant as well as the intracellular proteins are analyzed by SDS-PAGE followed by autoradiography (if radiolabeled).

Microsequencing of the amino acid sequence of the amino terminus of purified protein may be used to determine the amino terminal sequence of the produced protein.

Example 8: Expression of a Polypeptide in Mammalian Cells

10

15

20

25

30

The polypeptide of the present invention can be expressed in a mammalian cell. A typical mammalian expression vector contains a promoter element, which mediates the initiation of transcription of mRNA, a protein coding sequence, and signals required for the termination of transcription and polyadenylation of the transcript. Additional elements include enhancers, Kozak sequences and intervening sequences flanked by donor and acceptor sites for RNA splicing. Highly efficient transcription is achieved with the early and late promoters from SV40, the long terminal repeats (LTRs) from Retroviruses, e.g., RSV, HTLVI, HIVI and the early promoter of the cytomegalovirus (CMV). However, cellular elements can also be used (e.g., the human actin promoter).

Suitable expression vectors for use in practicing the present invention include, for example, vectors such as pSVL and pMSG (Pharmacia, Uppsala, Sweden), pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146), pBC12MI (ATCC 67109), pCMVSport 2.0, and pCMVSport 3.0. Mammalian host cells that could be used include, human Hela, 293, H9 and Jurkat cells, mouse NIH3T3 and C127 cells, Cos 1, Cos 7 and CV1, quail QC1-3 cells, mouse L cells and Chinese hamster ovary (CHO) cells.

Alternatively, the polypeptide can be expressed in stable cell lines containing the polynucleotide integrated into a chromosome. The co-transfection with a selectable marker such as dhfr, gpt, neomycin, hygromycin allows the identification and isolation of the transfected cells.

The transfected gene can also be amplified to express large amounts of the encoded protein. The DHFR (dihydrofolate reductase) marker is useful in developing cell lines that carry several hundred or even several thousand copies of the gene of interest. (See, e.g., Alt, F. W., et al., J. Biol. Chem. 253:1357-1370 (1978); Hamlin, J. L. and Ma, C., Biochem. et Biophys. Acta, 1097:107-143 (1990); Page, M. J. and Sydenham, M. A., Biotechnology 9:64-68 (1991).) Another useful selection marker is the enzyme glutamine synthase (GS) (Murphy et al., Biochem J. 227:277-279 (1991); Bebbington et al., Bio/Technology 10:169-175 (1992). Using these markers, the mammalian cells are grown in selective medium and the cells with the highest resistance are selected. These cell lines contain the amplified gene(s) integrated into a

10

15

20

25

30

chromosome. Chinese hamster ovary (CHO) and NSO cells are often used for the production of proteins.

Derivatives of the plasmid pSV2-dhfr (ATCC Accession No. 37146), the expression vectors pC4 (ATCC Accession No. 209646) and pC6 (ATCC Accession No.209647) contain the strong promoter (LTR) of the Rous Sarcoma Virus (Cullen et al., Molecular and Cellular Biology, 438-447 (March, 1985)) plus a fragment of the CMV-enhancer (Boshart et al., Cell 41:521-530 (1985).) Multiple cloning sites, e.g., with the restriction enzyme cleavage sites BamHI, XbaI and Asp718, facilitate the cloning of the gene of interest. The vectors also contain the 3' intron, the polyadenylation and termination signal of the rat preproinsulin gene, and the mouse DHFR gene under control of the SV40 early promoter.

Specifically, the plasmid pC6, for example, is digested with appropriate restriction enzymes and then dephosphorylated using calf intestinal phosphates by procedures known in the art. The vector is then isolated from a 1% agarose gel.

A polynucleotide of the present invention is amplified according to the protocol outlined in Example 1. If the naturally occurring signal sequence is used to produce the secreted protein, the vector does not need a second signal peptide.

Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

The amplified fragment is then digested with the same restriction enzyme and purified on a 1% agarose gel. The isolated fragment and the dephosphorylated vector are then ligated with T4 DNA ligase. *E. coli* HB101 or XL-1 Blue cells are then transformed and bacteria are identified that contain the fragment inserted into plasmid pC6 using, for instance, restriction enzyme analysis.

Chinese hamster ovary cells lacking an active DHFR gene is used for transfection. Five μg of the expression plasmid pC6 is cotransfected with 0.5 μg of the plasmid pSVneo using lipofectin (Felgner et al., supra). The plasmid pSV2-neo

10

20

25

30

contains a dominant selectable marker, the *neo* gene from Tn5 encoding an enzyme that confers resistance to a group of antibiotics including G418. The cells are seeded in alpha minus MEM supplemented with 1 mg/ml G418. After 2 days, the cells are trypsinized and seeded in hybridoma cloning plates (Greiner, Germany) in alpha minus MEM supplemented with 10, 25, or 50 ng/ml of metothrexate plus 1 mg/ml G418. After about 10-14 days single clones are trypsinized and then seeded in 6-well petri dishes or 10 ml flasks using different concentrations of methotrexate (50 nM, 100 nM, 200 nM, 400 nM, 800 nM). Clones growing at the highest concentrations of methotrexate are then transferred to new 6-well plates containing even higher concentrations of methotrexate (1 μ M, 2 μ M, 5 μ M, 10 mM, 20 mM). The same procedure is repeated until clones are obtained which grow at a concentration of 100 - 200 μ M. Expression of the desired gene product is analyzed, for instance, by SDS-PAGE and Western blot or by reversed phase HPLC analysis.

15 Example 9: Protein Fusions

The polypeptides of the present invention are preferably fused to other proteins. These fusion proteins can be used for a variety of applications. For example, fusion of the present polypeptides to His-tag, HA-tag, protein A, IgG domains, and maltose binding protein facilitates purification. (See Example 5; see also EP A 394,827; Traunecker, et al., Nature 331:84-86 (1988).) Similarly, fusion to IgG-1, IgG-3, and albumin increases the halflife time in vivo. Nuclear localization signals fused to the polypeptides of the present invention can target the protein to a specific subcellular localization, while covalent heterodimer or homodimers can increase or decrease the activity of a fusion protein. Fusion proteins can also create chimeric molecules having more than one function. Finally, fusion proteins can increase solubility and/or stability of the fused protein compared to the non-fused protein. All of the types of fusion proteins described above can be made by modifying the following protocol, which outlines the fusion of a polypeptide to an IgG molecule, or the protocol described in Example 5.

Briefly, the human Fc portion of the IgG molecule can be PCR amplified, using primers that span the 5' and 3' ends of the sequence described below. These

WO 99/47540 PCT/US99/05804

primers also should have convenient restriction enzyme sites that will facilitate cloning into an expression vector, preferably a mammalian expression vector.

For example, if pC4 (Accession No. 209646) is used, the human Fc portion can be ligated into the BamHI cloning site. Note that the 3' BamHI site should be destroyed. Next, the vector containing the human Fc portion is re-restricted with BamHI, linearizing the vector, and a polynucleotide of the present invention, isolated by the PCR protocol described in Example 1, is ligated into this BamHI site. Note that the polynucleotide is cloned without a stop codon, otherwise a fusion protein will not be produced.

If the naturally occurring signal sequence is used to produce the secreted protein, pC4 does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

Human IgG Fc region:

5

10

15

20

25

30

GGGATCCGGAGCCCAAATCTTCTGACAAAACTCACACATGCCCACCGTGC
CCAGCACCTGAATTCGAGGGTGCACCGTCAGTCTTCCTCTTCCCCCCAAAA
CCCAAGGACACCCTCATGATCTCCCGGACTCCTGAGGTCACATGCGTGGT
GGTGGACGTAAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGG
ACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTA
CAACAGCACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACT
GGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCA
ACCCCCATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAAC
CACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAG
GTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCAAGCGACATCGCCGT
GGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACTACAAGACCACGCCT
CCCGTGCTGGACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTG
GACAAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCCTCTCTCCGTGATGCA
TGAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCCCCTGTCTCCCGG
GTAAATGAGTGCGACGGCCGGGACTCTTAGAGGAT (SEQ ID NO:1)

5

10

15

20

25

30

Example 10: Production of an Antibody from a Polypeptide

The antibodies of the present invention can be prepared by a variety of methods. (See, Current Protocols, Chapter 2.) For example, cells expressing a polypeptide of the present invention is administered to an animal to induce the production of sera containing polyclonal antibodies. In a preferred method, a preparation of the secreted protein is prepared and purified to render it substantially free of natural contaminants. Such a preparation is then introduced into an animal in order to produce polyclonal antisera of greater specific activity.

In the most preferred method, the antibodies of the present invention are monoclonal antibodies (or protein binding fragments thereof). Such monoclonal antibodies can be prepared using hybridoma technology. (Köhler et al., Nature 256:495 (1975); Köhler et al., Eur. J. Immunol. 6:511 (1976); Köhler et al., Eur. J. Immunol. 6:292 (1976); Hammerling et al., in: Monoclonal Antibodies and T-Cell Hybridomas, Elsevier, N.Y., pp. 563-681 (1981).) In general, such procedures involve immunizing an animal (preferably a mouse) with polypeptide or, more preferably, with a secreted polypeptide-expressing cell. Such cells may be cultured in any suitable tissue culture medium; however, it is preferable to culture cells in Earle's modified Eagle's medium supplemented with 10% fetal bovine serum (inactivated at about 56°C), and supplemented with about 10 g/l of nonessential amino acids, about 1,000 U/ml of penicillin, and about 100 μg/ml of streptomycin.

The splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any suitable myeloma cell line may be employed in accordance with the present invention; however, it is preferable to employ the parent myeloma cell line (SP2O), available from the ATCC. After fusion, the resulting hybridoma cells are selectively maintained in HAT medium, and then cloned by limiting dilution as described by Wands et al. (Gastroenterology 80:225-232 (1981).) The hybridoma cells obtained through such a selection are then assayed to identify clones which secrete antibodies capable of binding the polypeptide.

Alternatively, additional antibodies capable of binding to the polypeptide can be produced in a two-step procedure using anti-idiotypic antibodies. Such a method makes use of the fact that antibodies are themselves antigens, and therefore, it is possible to obtain an antibody which binds to a second antibody. In accordance with this method, protein specific antibodies are used to immunize an animal, preferably a mouse. The splenocytes of such an animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody whose ability to bind to the protein-specific antibody can be blocked by the polypeptide. Such antibodies comprise anti-idiotypic antibodies to the protein-specific antibody and can be used to immunize an animal to induce formation of further protein-specific antibodies.

It will be appreciated that Fab and F(ab')2 and other fragments of the antibodies of the present invention may be used according to the methods disclosed herein. Such fragments are typically produced by proteolytic cleavage, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')2 fragments). Alternatively, secreted protein-binding fragments can be produced through the application of recombinant DNA technology or through synthetic chemistry.

For in vivo use of antibodies in humans, it may be preferable to use "humanized" chimeric monoclonal antibodies. Such antibodies can be produced using genetic constructs derived from hybridoma cells producing the monoclonal antibodies described above. Methods for producing chimeric antibodies are known in the art. (See, for review, Morrison, Science 229:1202 (1985); Oi et al., BioTechniques 4:214 (1986); Cabilly et al., U.S. Patent No. 4,816,567; Taniguchi et al., EP 171496; Morrison et al., EP 173494; Neuberger et al., WO 8601533; Robinson et al., WO 8702671; Boulianne et al., Nature 312:643 (1984); Neuberger et al., Nature 314:268 (1985).)

25

30

20

5

10

15

Example 11: Production Of Secreted Protein For High-Throughput Screening Assays

The following protocol produces a supernatant containing a polypeptide to be tested. This supernatant can then be used in the Screening Assays described in Examples 13-20.

WO 99/47540 PCT/US99/05804

First, dilute Poly-D-Lysine (644 587 Boehringer-Mannheim) stock solution (1mg/ml in PBS) 1:20 in PBS (w/o calcium or magnesium 17-516F Biowhittaker) for a working solution of 50ug/ml. Add 200 ul of this solution to each well (24 well plates) and incubate at RT for 20 minutes. Be sure to distribute the solution over each well (note: a 12-channel pipetter may be used with tips on every other channel). Aspirate off the Poly-D-Lysine solution and rinse with 1ml PBS (Phosphate Buffered Saline). The PBS should remain in the well until just prior to plating the cells and plates may be poly-lysine coated in advance for up to two weeks.

Plate 293T cells (do not carry cells past P+20) at 2 x 10⁵ cells/well in .5ml DMEM(Dulbecco's Modified Eagle Medium)(with 4.5 G/L glucose and L-glutamine (12-604F Biowhittaker))/10% heat inactivated FBS(14-503F Biowhittaker)/1x Penstrep(17-602E Biowhittaker). Let the cells grow overnight.

10

15

20

25

30

The next day, mix together in a sterile solution basin: 300 ul Lipofectamine (18324-012 Gibco/BRL) and 5ml Optimem I (31985070 Gibco/BRL)/96-well plate.

With a small volume multi-channel pipetter, aliquot approximately 2ug of an expression vector containing a polynucleotide insert, produced by the methods described in Examples 8 or 9, into an appropriately labeled 96-well round bottom plate. With a multi-channel pipetter, add 50ul of the Lipofectamine/Optimem I mixture to each well. Pipette up and down gently to mix. Incubate at RT 15-45 minutes. After about 20 minutes, use a multi-channel pipetter to add 150ul Optimem I to each well. As a control, one plate of vector DNA lacking an insert should be transfected with each set of transfections.

Preferably, the transfection should be performed by tag-teaming the following tasks. By tag-teaming, hands on time is cut in half, and the cells do not spend too much time on PBS. First, person A aspirates off the media from four 24-well plates of cells, and then person B rinses each well with .5-1ml PBS. Person A then aspirates off PBS rinse, and person B, using a12-channel pipetter with tips on every other channel, adds the 200ul of DNA/Lipofectamine/Optimem I complex to the odd wells first, then to the even wells, to each row on the 24-well plates. Incubate at 37°C for 6 hours.

While cells are incubating, prepare appropriate media, either 1%BSA in DMEM with 1x penstrep, or CHO-5 media (116.6 mg/L of CaCl2 (anhyd); 0.00130 $mg/L CuSO_4-5H_2O$; 0.050 mg/L of $Fe(NO_3)_3-9H_2O$; 0.417 mg/L of $FeSO_4-7H_2O$; 311.80 mg/L of Kcl; 28.64 mg/L of MgCl₂; 48.84 mg/L of MgSO₄; 6995.50 mg/L of NaCl; 2400.0 mg/L of NaHCO₃; 62.50 mg/L of NaH₂PO₄-H₂0; 71.02 mg/L of 5 Na, HPO4; .4320 mg/L of ZnSO₄-7H₂O; .002 mg/L of Arachidonic Acid; 1.022 mg/L of Cholesterol; .070 mg/L of DL-alpha-Tocopherol-Acetate; 0.0520 mg/L of Linoleic Acid; 0.010 mg/L of Linolenic Acid; 0.010 mg/L of Myristic Acid; 0.010 mg/L of Oleic Acid; 0.010 mg/L of Palmitric Acid; 0.010 mg/L of Palmitic Acid; 100 mg/L of Pluronic F-68; 0.010 mg/L of Stearic Acid; 2.20 mg/L of Tween 80; 4551 mg/L of D-10 Glucose; 130.85 mg/ml of L- Alanine; 147.50 mg/ml of L-Arginine-HCL; 7.50 mg/ml of L-Asparagine-H₂0; 6.65 mg/ml of L-Aspartic Acid; 29.56 mg/ml of L-Cystine-2HCL-H₂0; 31.29 mg/ml of L-Cystine-2HCL; 7.35 mg/ml of L-Glutamic Acid; 365.0 mg/ml of L-Glutamine; 18.75 mg/ml of Glycine; 52.48 mg/ml of L-Histidine-HCL-H₂0; 106.97 mg/ml of L-Isoleucine; 111.45 mg/ml of L-Leucine; 163.75 mg/ml of L-15 Lysine HCL; 32.34 mg/ml of L-Methionine; 68.48 mg/ml of L-Phenylalainine; 40.0 mg/ml of L-Proline; 26.25 mg/ml of L-Serine; 101.05 mg/ml of L-Threonine; 19.22 mg/ml of L-Tryptophan; 91.79 mg/ml of L-Tryrosine-2Na-2H₂0; 99.65 mg/ml of L-Valine; 0.0035 mg/L of Biotin; 3.24 mg/L of D-Ca Pantothenate; 11.78 mg/L of Choline Chloride; 4.65 mg/L of Folic Acid; 15.60 mg/L of i-Inositol; 3.02 mg/L of 20 Niacinamide; 3.00 mg/L of Pyridoxal HCL; 0.031 mg/L of Pyridoxine HCL; 0.319 mg/L of Riboflavin; 3.17 mg/L of Thiamine HCL; 0.365 mg/L of Thymidine; and 0.680 mg/L of Vitamin B₁₂; 25 mM of HEPES Buffer; 2.39 mg/L of Na Hypoxanthine; 0.105 mg/L of Lipoic Acid; 0.081 mg/L of Sodium Putrescine-2HCL; 25 55.0 mg/L of Sodium Pyruvate; 0.0067 mg/L of Sodium Selenite; 20uM of Ethanolamine; 0.122 mg/L of Ferric Citrate; 41.70 mg/L of Methyl-B-Cyclodextrin complexed with Linoleic Acid; 33.33 mg/L of Methyl-B-Cyclodextrin complexed with Oleic Acid; and 10 mg/L of Methyl-B-Cyclodextrin complexed with Retinal) with 2mm glutamine and 1x penstrep. (BSA (81-068-3 Bayer) 100gm dissolved in 1L DMEM for a 10% BSA stock solution). Filter the media and collect 50 ul for 30 endotoxin assay in 15ml polystyrene conical.

PCT/US99/05804

The transfection reaction is terminated, preferably by tag-teaming, at the end of the incubation period. Person A aspirates off the transfection media, while person B adds 1.5ml appropriate media to each well. Incubate at 37°C for 45 or 72 hours depending on the media used: 1%BSA for 45 hours or CHO-5 for 72 hours.

On day four, using a 300ul multichannel pipetter, aliquot 600ul in one 1ml deep well plate and the remaining supernatant into a 2ml deep well. The supernatants from each well can then be used in the assays described in Examples 13-20.

It is specifically understood that when activity is obtained in any of the assays described below using a supernatant, the activity originates from either the polypeptide directly (e.g., as a secreted protein) or by the polypeptide inducing expression of other proteins, which are then secreted into the supernatant. Thus, the invention further provides a method of identifying the protein in the supernatant characterized by an activity in a particular assay.

15 Example 12: Construction of GAS Reporter Construct

5

10

20

25

30

One signal transduction pathway involved in the differentiation and proliferation of cells is called the Jaks-STATs pathway. Activated proteins in the Jaks-STATs pathway bind to gamma activation site "GAS" elements or interferonsensitive responsive element ("ISRE"), located in the promoter of many genes. The binding of a protein to these elements alter the expression of the associated gene.

GAS and ISRE elements are recognized by a class of transcription factors called Signal Transducers and Activators of Transcription, or "STATs." There are six members of the STATs family. Stat1 and Stat3 are present in many cell types, as is Stat2 (as response to IFN-alpha is widespread). Stat4 is more restricted and is not in many cell types though it has been found in T helper class I, cells after treatment with IL-12. Stat5 was originally called mammary growth factor, but has been found at higher concentrations in other cells including myeloid cells. It can be activated in tissue culture cells by many cytokines.

The STATs are activated to translocate from the cytoplasm to the nucleus upon tyrosine phosphorylation by a set of kinases known as the Janus Kinase ("Jaks") family. Jaks represent a distinct family of soluble tyrosine kinases and include Tyk2,

Jak1, Jak2, and Jak3. These kinases display significant sequence similarity and are generally catalytically inactive in resting cells.

The Jaks are activated by a wide range of receptors summarized in the Table below. (Adapted from review by Schidler and Darnell, Ann. Rev. Biochem. 64:621-51 (1995).) A cytokine receptor family, capable of activating Jaks, is divided into two groups: (a) Class 1 includes receptors for IL-2, IL-3, IL-4, IL-6, IL-7, IL-9, IL-11, IL-12, IL-15, Epo, PRL, GH, G-CSF, GM-CSF, LIF, CNTF, and thrombopoietin; and (b) Class 2 includes IFN-a, IFN-g, and IL-10. The Class 1 receptors share a conserved cysteine motif (a set of four conserved cysteines and one tryptophan) and a WSXWS motif (a membrane proximal region encoding Trp-Ser-Xxx-Trp-Ser (SEQ ID NO:2)).

5

10

Thus, on binding of a ligand to a receptor, Jaks are activated, which in turn activate STATs, which then translocate and bind to GAS elements. This entire process is encompassed in the Jaks-STATs signal transduction pathway.

Therefore, activation of the Jaks-STATs pathway, reflected by the binding of
the GAS or the ISRE element, can be used to indicate proteins involved in the
proliferation and differentiation of cells. For example, growth factors and cytokines
are known to activate the Jaks-STATs pathway. (See Table below.) Thus, by using
GAS elements linked to reporter molecules, activators of the Jaks-STATs pathway
can be identified.

Ligand	tyk2	<u>JAKs</u> <u>Jak l</u>	Jak2	Jak3	STATS	GAS(elements) or ISRE
IFN family IFN-a/B IFN-g II-10	+	+ + ?	- + ?		1,2,3 1 1,3	ISRE GAS (IRF1>Lys6>IFP)
gp130 family IL-6 (Pleiotrophic) Il-11(Pleiotrophic) OnM(Pleiotrophic) LIF(Pleiotrophic) CNTF(Pleiotrophic) G-CSF(Pleiotrophic) IL-12(Pleiotrophic)	+ ? ? ? -/+ ?	+ + + + +	+ ? + + + + ? +	? ? ? ? ? ? +	1,3 1,3 1,3 1,3 1,3 1,3	GAS (IRF1>Lys6>IFP)
g-C family IL-2 (lymphocytes) IL-4 (lymph/myeloid) IL-7 (lymphocytes) IL-9 (lymphocytes) IL-13 (lymphocyte) IL-15	- - - - ?	+ + + + + + +	- - - ? ?	+ + + + ?	1,3,5 6 5 5 6 5	GAS GAS (IRF1 = IFP >>Ly6)(IgH) GAS GAS GAS GAS
gp140 family IL-3 (myeloid) IL-5 (myeloid) GM-CSF (myeloid)	- - -	:	+ + +	:	5 5 5	GAS (IRF1>IFP>>Ly6) GAS GAS
Growth hormone fami GH PRL EPO	l <u>y</u> ? ? ?	- +/- -	+ + +	-	5 1,3,5 5	GAS(B-CAS>IRF1=IFP>>Ly6)
Receptor Tyrosine Kir EGF PDGF CSF-1	nases ? ? ?	+ + +	+ + +	-	1,3 1,3 1,3	GAS (IRF1) GAS (not IRF1)

PCT/US99/05804

To construct a synthetic GAS containing promoter element, which is used in the Biological Assays described in Examples 13-14, a PCR based strategy is employed to generate a GAS-SV40 promoter sequence. The 5' primer contains four tandem copies of the GAS binding site found in the IRF1 promoter and previously demonstrated to bind STATs upon induction with a range of cytokines (Rothman et al., Immunity 1:457-468 (1994).), although other GAS or ISRE elements can be used instead. The 5' primer also contains 18bp of sequence complementary to the SV40 early promoter sequence and is flanked with an XhoI site. The sequence of the 5' primer is:

10 5':GCGCCTCGAGATTTCCCCGAAATCTAGATTTCCCCGAAATGATTTCCCC GAAATGATTTCCCCGAAATATCTGCCATCTCAATTAG:3' (SEQ ID NO:3)

The downstream primer is complementary to the SV40 promoter and is flanked with a Hind III site: 5':GCGGCAAGCTTTTTGCAAAGCCTAGGC:3' (SEQ ID NO:4)

15

30

PCR amplification is performed using the SV40 promoter template present in the B-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI/Hind III and subcloned into BLSK2-. (Stratagene.) Sequencing with forward and reverse primers confirms that the insert contains the following sequence:

5':CTCGAGATTTCCCCGAAATCTAGATTTCCCCGAAATGATTTCCCCGAAA TGATTTCCCCGAAATATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCG CCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCT CCGCCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCC TCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTTGGAGGCCT 25 AGGCTTTTGCAAAAAAGCTT:3' (SEQ ID NO:5)

With this GAS promoter element linked to the SV40 promoter, a GAS:SEAP2 reporter construct is next engineered. Here, the reporter molecule is a secreted alkaline phosphatase, or "SEAP." Clearly, however, any reporter molecule can be instead of SEAP, in this or in any of the other Examples. Well known reporter molecules that can be used instead of SEAP include chloramphenicol

10

15

20

acetyltransferase (CAT), luciferase, alkaline phosphatase, B-galactosidase, green fluorescent protein (GFP), or any protein detectable by an antibody.

The above sequence confirmed synthetic GAS-SV40 promoter element is subcloned into the pSEAP-Promoter vector obtained from Clontech using HindIII and XhoI, effectively replacing the SV40 promoter with the amplified GAS:SV40 promoter element, to create the GAS-SEAP vector. However, this vector does not contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

Thus, in order to generate mammalian stable cell lines expressing the GAS-SEAP reporter, the GAS-SEAP cassette is removed from the GAS-SEAP vector using SalI and NotI, and inserted into a backbone vector containing the neomycin resistance gene, such as pGFP-1 (Clontech), using these restriction sites in the multiple cloning site, to create the GAS-SEAP/Neo vector. Once this vector is transfected into mammalian cells, this vector can then be used as a reporter molecule for GAS binding as described in Examples 13-14.

Other constructs can be made using the above description and replacing GAS with a different promoter sequence. For example, construction of reporter molecules containing NFK-B and EGR promoter sequences are described in Examples 15 and 16. However, many other promoters can be substituted using the protocols described in these Examples. For instance, SRE, IL-2, NFAT, or Osteocalcin promoters can be substituted, alone or in combination (e.g., GAS/NF-KB/EGR, GAS/NF-KB, Il-2/NFAT, or NF-KB/GAS). Similarly, other cell lines can be used to test reporter construct activity, such as HELA (epithelial), HUVEC (endothelial), Reh (B-cell), Saos-2 (osteoblast), HUVAC (aortic), or Cardiomyocyte.

25

30

Example 13: High-Throughput Screening Assay for T-cell Activity.

The following protocol is used to assess T-cell activity by identifying factors, such as growth factors and cytokines, that may proliferate or differentiate T-cells. T-cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the Jaks-STATS signal transduction pathway. The T-cell used in this assay is Jurkat T-cells

10

15

20

25

30

(ATCC Accession No. TIB-152), although Molt-3 cells (ATCC Accession No. CRL-1552) and Molt-4 cells (ATCC Accession No. CRL-1582) cells can also be used.

Jurkat T-cells are lymphoblastic CD4+ Th1 helper cells. In order to generate stable cell lines, approximately 2 million Jurkat cells are transfected with the GAS-SEAP/neo vector using DMRIE-C (Life Technologies)(transfection procedure described below). The transfected cells are seeded to a density of approximately 20,000 cells per well and transfectants resistant to 1 mg/ml genticin selected. Resistant colonies are expanded and then tested for their response to increasing concentrations of interferon gamma. The dose response of a selected clone is demonstrated.

Specifically, the following protocol will yield sufficient cells for 75 wells containing 200 ul of cells. Thus, it is either scaled up, or performed in multiple to generate sufficient cells for multiple 96 well plates. Jurkat cells are maintained in RPMI + 10% serum with 1%Pen-Strep. Combine 2.5 mls of OPTI-MEM (Life Technologies) with 10 ug of plasmid DNA in a T25 flask. Add 2.5 ml OPTI-MEM containing 50 ul of DMRIE-C and incubate at room temperature for 15-45 mins.

During the incubation period, count cell concentration, spin down the required number of cells (10⁷ per transfection), and resuspend in OPTI-MEM to a final concentration of 10⁷ cells/ml. Then add 1ml of 1 x 10⁷ cells in OPTI-MEM to T25 flask and incubate at 37°C for 6 hrs. After the incubation, add 10 ml of RPMI + 15% serum.

The Jurkat: GAS-SEAP stable reporter lines are maintained in RPMI + 10% serum, 1 mg/ml Genticin, and 1% Pen-Strep. These cells are treated with supernatants containing a polypeptide as produced by the protocol described in Example 11.

On the day of treatment with the supernatant, the cells should be washed and resuspended in fresh RPMI + 10% serum to a density of 500,000 cells per ml. The exact number of cells required will depend on the number of supernatants being screened. For one 96 well plate, approximately 10 million cells (for 10 plates, 100 million cells) are required.

10

15

25

Transfer the cells to a triangular reservoir boat, in order to dispense the cells into a 96 well dish, using a 12 channel pipette. Using a 12 channel pipette, transfer 200 ul of cells into each well (therefore adding 100, 000 cells per well).

After all the plates have been seeded, 50 ul of the supernatants are transferred directly from the 96 well plate containing the supernatants into each well using a 12 channel pipette. In addition, a dose of exogenous interferon gamma (0.1, 1.0, 10 ng) is added to wells H9, H10, and H11 to serve as additional positive controls for the assay.

The 96 well dishes containing Jurkat cells treated with supernatants are placed in an incubator for 48 hrs (note: this time is variable between 48-72 hrs). 35 ul samples from each well are then transferred to an opaque 96 well plate using a 12 channel pipette. The opaque plates should be covered (using sellophene covers) and stored at -20°C until SEAP assays are performed according to Example 17. The plates containing the remaining treated cells are placed at 4°C and serve as a source of material for repeating the assay on a specific well if desired.

As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate Jurkat T cells. Over 30 fold induction is typically observed in the positive control wells.

The above protocol may be used in the generation of both transient, as well as,
stable transfected cells, which would be apparent to those of skill in the art.

Example 14: High-Throughput Screening Assay Identifying Myeloid Activity

The following protocol is used to assess myeloid activity by identifying factors, such as growth factors and cytokines, that may proliferate or differentiate myeloid cells. Myeloid cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the Jaks-STATS signal transduction pathway. The myeloid cell used in this assay is U937, a pre-monocyte cell line, although TF-1, HL60, or KG1 can be used.

To transiently transfect U937 cells with the GAS/SEAP/Neo construct produced in Example 12, a DEAE-Dextran method (Kharbanda et. al., 1994, Cell

Growth & Differentiation, 5:259-265) is used. First, harvest 2x10e⁷ U937 cells and wash with PBS. The U937 cells are usually grown in RPMI 1640 medium containing 10% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 mg/ml streptomycin.

Next, suspend the cells in 1 ml of 20 mM Tris-HCl (pH 7.4) buffer containing 0.5 mg/ml DEAE-Dextran, 8 ug GAS-SEAP2 plasmid DNA, 140 mM NaCl, 5 mM KCl, 375 uM Na₂HPO₄.7H₂O, 1 mM MgCl₂, and 675 uM CaCl₂. Incubate at 37°C for 45 min.

Wash the cells with RPMI 1640 medium containing 10% FBS and then resuspend in 10 ml complete medium and incubate at 37°C for 36 hr.

5

10

15

20

The GAS-SEAP/U937 stable cells are obtained by growing the cells in 400 ug/ml G418. The G418-free medium is used for routine growth but every one to two months, the cells should be re-grown in 400 ug/ml G418 for couple of passages.

These cells are tested by harvesting $1x10^8$ cells (this is enough for ten 96-well plates assay) and wash with PBS. Suspend the cells in 200 ml above described growth medium, with a final density of $5x10^5$ cells/ml. Plate 200 ul cells per well in the 96-well plate (or $1x10^5$ cells/well).

Add 50 ul of the supernatant prepared by the protocol described in Example 11. Incubate at 37°C for 48 to 72 hr. As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate U937 cells. Over 30 fold induction is typically observed in the positive control wells. SEAP assay the supernatant according to the protocol described in Example 17.

Example 15: High-Throughput Screening Assay Identifying Neuronal Activity.

When cells undergo differentiation and proliferation, a group of genes are activated through many different signal transduction pathways. One of these genes, EGR1 (early growth response gene 1), is induced in various tissues and cell types upon activation. The promoter of EGR1 is responsible for such induction. Using the EGR1 promoter linked to reporter molecules, activation of cells can be assessed.

Particularly, the following protocol is used to assess neuronal activity in PC12 cell lines. PC12 cells (rat phenochromocytoma cells) are known to proliferate and/or differentiate by activation with a number of mitogens, such as TPA (tetradecanoyl phorbol acetate), NGF (nerve growth factor), and EGF (epidermal growth factor).

The EGR1 gene expression is activated during this treatment. Thus, by stably transfecting PC12 cells with a construct containing an EGR promoter linked to SEAP reporter, activation of PC12 cells can be assessed.

The EGR/SEAP reporter construct can be assembled by the following protocol. The EGR-1 promoter sequence (-633 to +1)(Sakamoto K et al., Oncogene 6:867-871 (1991)) can be PCR amplified from human genomic DNA using the following primers:

10

15

20

25

30

- 5' GCGCTCGAGGGATGACAGCGATAGAACCCCGG -3' (SEQ ID NO:6)
- 5' GCGAAGCTTCGCGACTCCCCGGATCCGCCTC-3' (SEO ID NO:7)

Using the GAS:SEAP/Neo vector produced in Example 12, EGR1 amplified product can then be inserted into this vector. Linearize the GAS:SEAP/Neo vector using restriction enzymes XhoI/HindIII, removing the GAS/SV40 stuffer. Restrict the EGR1 amplified product with these same enzymes. Ligate the vector and the EGR1 promoter.

To prepare 96 well-plates for cell culture, two mls of a coating solution (1:30 dilution of collagen type I (Upstate Biotech Inc. Cat#08-115) in 30% ethanol (filter sterilized)) is added per one 10 cm plate or 50 ml per well of the 96-well plate, and allowed to air dry for 2 hr.

PC12 cells are routinely grown in RPMI-1640 medium (Bio Whittaker) containing 10% horse serum (JRH BIOSCIENCES, Cat. # 12449-78P), 5% heatinactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 ug/ml streptomycin on a precoated 10 cm tissue culture dish. One to four split is done every three to four days. Cells are removed from the plates by scraping and resuspended with pipetting up and down for more than 15 times.

Transfect the EGR/SEAP/Neo construct into PC12 using the Lipofectamine protocol described in Example 11. EGR-SEAP/PC12 stable cells are obtained by growing the cells in 300 ug/ml G418. The G418-free medium is used for routine

10

15

20

25

30

growth but every one to two months, the cells should be re-grown in 300 ug/ml G418 for couple of passages.

To assay for neuronal activity, a 10 cm plate with cells around 70 to 80% confluent is screened by removing the old medium. Wash the cells once with PBS (Phosphate buffered saline). Then starve the cells in low serum medium (RPMI-1640 containing 1% horse serum and 0.5% FBS with antibiotics) overnight.

The next morning, remove the medium and wash the cells with PBS. Scrape off the cells from the plate, suspend the cells well in 2 ml low serum medium. Count the cell number and add more low serum medium to reach final cell density as $5x10^5$ cells/ml.

Add 200 ul of the cell suspension to each well of 96-well plate (equivalent to $1x10^5$ cells/well). Add 50 ul supernatant produced by Example 11, 37°C for 48 to 72 hr. As a positive control, a growth factor known to activate PC12 cells through EGR can be used, such as 50 ng/ul of Neuronal Growth Factor (NGF). Over fifty-fold induction of SEAP is typically seen in the positive control wells. SEAP assay the supernatant according to Example 17.

Example 16: High-Throughput Screening Assay for T-cell Activity

NF-κB (Nuclear Factor κB) is a transcription factor activated by a wide variety of agents including the inflammatory cytokines IL-1 and TNF, CD30 and CD40, lymphotoxin-alpha and lymphotoxin-beta, by exposure to LPS or thrombin, and by expression of certain viral gene products. As a transcription factor, NF-κB regulates the expression of genes involved in immune cell activation, control of apoptosis (NF- κB appears to shield cells from apoptosis), B and T-cell development, anti-viral and antimicrobial responses, and multiple stress responses.

In non-stimulated conditions, NF- κB is retained in the cytoplasm with I-κB (Inhibitor κB). However, upon stimulation, I- κB is phosphorylated and degraded, causing NF- κB to shuttle to the nucleus, thereby activating transcription of target genes. Target genes activated by NF- κB include IL-2, IL-6, GM-CSF, ICAM-1 and class 1 MHC.

10

20

25

Due to its central role and ability to respond to a range of stimuli, reporter constructs utilizing the NF-kB promoter element are used to screen the supernatants produced in Example 11. Activators or inhibitors of NF-kB would be useful in treating diseases. For example, inhibitors of NF-kB could be used to treat those diseases related to the acute or chronic activation of NF-kB, such as rheumatoid arthritis.

The downstream primer is complementary to the 3' end of the SV40 promoter and is flanked with a Hind III site:

15 5':GCGGCAAGCTTTTTGCAAAGCCTAGGC:3' (SEQ ID NO:4)

PCR amplification is performed using the SV40 promoter template present in the pB-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI and Hind III and subcloned into BLSK2-. (Stratagene)

Sequencing with the T7 and T3 primers confirms the insert contains the following sequence:

5':CTCGAGGGGACTTTCCCGGGGACTTTCCGGGACTTTCC
ATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCC
ATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGA
CTAATTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTA
TTCCAGAAGTAGTGAGGAGGCCTTTTTTGGAGGCCTAGGCTTTTGCAAAAA
GCTT:3' (SEQ ID NO:10)

Next, replace the SV40 minimal promoter element present in the pSEAP2-30 promoter plasmid (Clontech) with this NF-κB/SV40 fragment using XhoI and

15

20

25

30

HindIII. However, this vector does not contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

In order to generate stable mammalian cell lines, the NF-kB/SV40/SEAP cassette is removed from the above NF-kB/SEAP vector using restriction enzymes Sall and NotI, and inserted into a vector containing neomycin resistance. Particularly, the NF-kB/SV40/SEAP cassette was inserted into pGFP-1 (Clontech), replacing the GFP gene, after restricting pGFP-1 with Sall and NotI.

Once NF-kB/SV40/SEAP/Neo vector is created, stable Jurkat T-cells are created and maintained according to the protocol described in Example 13. Similarly, the method for assaying supernatants with these stable Jurkat T-cells is also described in Example 13. As a positive control, exogenous TNF alpha (0.1,1, 10 ng) is added to wells H9, H10, and H11, with a 5-10 fold activation typically observed.

Example 17: Assay for SEAP Activity

As a reporter molecule for the assays described in Examples 13-16, SEAP activity is assayed using the Tropix Phospho-light Kit (Cat. BP-400) according to the following general procedure. The Tropix Phospho-light Kit supplies the Dilution, Assay, and Reaction Buffers used below.

Prime a dispenser with the 2.5x Dilution Buffer and dispense $15 \,\mu l$ of 2.5x dilution buffer into Optiplates containing $35 \,\mu l$ of a supernatant. Seal the plates with a plastic sealer and incubate at 65°C for 30 min. Separate the Optiplates to avoid uneven heating.

Cool the samples to room temperature for 15 minutes. Empty the dispenser and prime with the Assay Buffer. Add 50 μ l Assay Buffer and incubate at room temperature 5 min. Empty the dispenser and prime with the Reaction Buffer (see the table below). Add 50 μ l Reaction Buffer and incubate at room temperature for 20 minutes. Since the intensity of the chemiluminescent signal is time dependent, and it takes about 10 minutes to read 5 plates on luminometer, one should treat 5 plates at each time and start the second set 10 minutes later.

Read the relative light unit in the luminometer. Set H12 as blank, and print the results. An increase in chemiluminescence indicates reporter activity.

Reaction Buffer Formulat	ion:
--------------------------	------

# of plates	Rxn buffer diluent (ml)	CSPD (ml)
10	60	3
11	65	3.25
12	70	3.5
13	75	3.75
14	80	4
15	85	4.25
16	90	4.5
17	95	4.75
18	· 100	5
19	105	5.25
20	110	5.5
21	115	5.75
22	120	6
23	125	6.25
24	130	6.5
25	135	6.75
26	140	7
27	145	7.25
28	150	7.5
29	155	7.75
30	160	8
31	165	8.25
32	170	8.5
33	175	8.75
34	180	9
35	185	9.25
36	190	9.5
37	195	9.75
38	200	10
39	205	10.25
40	210	10.5
41	215	10.75
42	220	11
43	225	11.25
44	230	11.5
45	235	11.75
46	240	12
47	245	12.25
48	250	12.5
49	255	12.75
50	260	13

Example 18: High-Throughput Screening Assay Identifying Changes in Small

5 Molecule Concentration and Membrane Permeability

Binding of a ligand to a receptor is known to alter intracellular levels of small molecules, such as calcium, potassium, sodium, and pH, as well as alter membrane potential. These alterations can be measured in an assay to identify supernatants

10

15

20

25

30

which bind to receptors of a particular cell. Although the following protocol describes an assay for calcium, this protocol can easily be modified to detect changes in potassium, sodium, pH, membrane potential, or any other small molecule which is detectable by a fluorescent probe.

The following assay uses Fluorometric Imaging Plate Reader ("FLIPR") to measure changes in fluorescent molecules (Molecular Probes) that bind small molecules. Clearly, any fluorescent molecule detecting a small molecule can be used instead of the calcium fluorescent molecule, fluo-4 (Molecular Probes, Inc.; catalog no. F-14202), used here.

For adherent cells, seed the cells at 10,000 -20,000 cells/well in a Co-star black 96-well plate with clear bottom. The plate is incubated in a CO₂ incubator for 20 hours. The adherent cells are washed two times in Biotek washer with 200 ul of HBSS (Hank's Balanced Salt Solution) leaving 100 ul of buffer after the final wash.

A stock solution of 1 mg/ml fluo-4 is made in 10% pluronic acid DMSO. To load the cells with fluo-4, 50 ul of 12 ug/ml fluo-4 is added to each well. The plate is incubated at 37° C in a CO_2 incubator for 60 min. The plate is washed four times in the Biotek washer with HBSS leaving 100 ul of buffer.

For non-adherent cells, the cells are spun down from culture media. Cells are re-suspended to 2-5x10⁶ cells/ml with HBSS in a 50-ml conical tube. 4 ul of 1 mg/ml fluo-4 solution in 10% pluronic acid DMSO is added to each ml of cell suspension. The tube is then placed in a 37°C water bath for 30-60 min. The cells are washed twice with HBSS, resuspended to 1x10⁶ cells/ml, and dispensed into a microplate, 100 ul/well. The plate is centrifuged at 1000 rpm for 5 min. The plate is then washed once in Denley CellWash with 200 ul, followed by an aspiration step to 100 ul final volume.

For a non-cell based assay, each well contains a fluorescent molecule, such as fluo-4. The supernatant is added to the well, and a change in fluorescence is detected.

To measure the fluorescence of intracellular calcium, the FLIPR is set for the following parameters: (1) System gain is 300-800 mW; (2) Exposure time is 0.4 second; (3) Camera F/stop is F/2; (4) Excitation is 488 nm; (5) Emission is 530 nm;

15

20

25

30

and (6) Sample addition is 50 ul. Increased emission at 530 nm indicates an extracellular signaling event which has resulted in an increase in the intracellular Ca++ concentration.

5 Example 19: High-Throughput Screening Assay Identifying Tyrosine Kinase Activity

The Protein Tyrosine Kinases (PTK) represent a diverse group of transmembrane and cytoplasmic kinases. Within the Receptor Protein Tyrosine Kinase RPTK) group are receptors for a range of mitogenic and metabolic growth factors including the PDGF, FGF, EGF, NGF, HGF and Insulin receptor subfamilies. In addition there are a large family of RPTKs for which the corresponding ligand is unknown. Ligands for RPTKs include mainly secreted small proteins, but also membrane-bound and extracellular matrix proteins.

Activation of RPTK by ligands involves ligand-mediated receptor dimerization, resulting in transphosphorylation of the receptor subunits and activation of the cytoplasmic tyrosine kinases. The cytoplasmic tyrosine kinases include receptor associated tyrosine kinases of the src-family (e.g., src, yes, lck, lyn, fyn) and non-receptor linked and cytosolic protein tyrosine kinases, such as the Jak family, members of which mediate signal transduction triggered by the cytokine superfamily of receptors (e.g., the Interleukins, Interferons, GM-CSF, and Leptin).

Because of the wide range of known factors capable of stimulating tyrosine kinase activity, the identification of novel human secreted proteins capable of activating tyrosine kinase signal transduction pathways are of interest. Therefore, the following protocol is designed to identify those novel human secreted proteins capable of activating the tyrosine kinase signal transduction pathways.

Seed target cells (e.g., primary keratinocytes) at a density of approximately 25,000 cells per well in a 96 well Loprodyne Silent Screen Plates purchased from Nalge Nunc (Naperville, IL). The plates are sterilized with two 30 minute rinses with 100% ethanol, rinsed with water and dried overnight. Some plates are coated for 2 hr with 100 ml of cell culture grade type I collagen (50 mg/ml), gelatin (2%) or polylysine (50 mg/ml), all of which can be purchased from Sigma Chemicals (St.

25

30

Louis, MO) or 10% Matrigel purchased from Becton Dickinson (Bedford,MA), or calf serum, rinsed with PBS and stored at 4°C. Cell growth on these plates is assayed by seeding 5,000 cells/well in growth medium and indirect quantitation of cell number through use of alamarBlue as described by the manufacturer Alamar Biosciences, Inc. (Sacramento, CA) after 48 hr. Falcon plate covers #3071 from Becton Dickinson (Bedford,MA) are used to cover the Loprodyne Silent Screen Plates. Falcon Microtest III cell culture plates can also be used in some proliferation experiments.

To prepare extracts, A431 cells are seeded onto the nylon membranes of Loprodyne plates (20,000/200ml/well) and cultured overnight in complete medium. 10 Cells are quiesced by incubation in serum-free basal medium for 24 hr. After 5-20 minutes treatment with EGF (60ng/ml) or 50 ul of the supernatant produced in Example 11, the medium was removed and 100 ml of extraction buffer ((20 mM HEPES pH 7.5, 0.15 M NaCl, 1% Triton X-100, 0.1% SDS, 2 mM Na3VO4, 2 mM Na4P2O7 and a cocktail of protease inhibitors (# 1836170) obtained from 15 Boeheringer Mannheim (Indianapolis, IN) is added to each well and the plate is shaken on a rotating shaker for 5 minutes at 4°C. The plate is then placed in a vacuum transfer manifold and the extract filtered through the 0.45 mm membrane bottoms of each well using house vacuum. Extracts are collected in a 96-well 20 catch/assay plate in the bottom of the vacuum manifold and immediately placed on ice. To obtain extracts clarified by centrifugation, the content of each well, after detergent solubilization for 5 minutes, is removed and centrifuged for 15 minutes at 4°C at 16,000 x g.

Test the filtered extracts for levels of tyrosine kinase activity. Although many methods of detecting tyrosine kinase activity are known, one method is described here.

Generally, the tyrosine kinase activity of a supernatant is evaluated by determining its ability to phosphorylate a tyrosine residue on a specific substrate (a biotinylated peptide). Biotinylated peptides that can be used for this purpose include PSK1 (corresponding to amino acids 6-20 of the cell division kinase cdc2-p34) and

15

20

25

30

PSK2 (corresponding to amino acids 1-17 of gastrin). Both peptides are substrates for a range of tyrosine kinases and are available from Boehringer Mannheim.

The tyrosine kinase reaction is set up by adding the following components in order. First, add 10ul of 5uM Biotinylated Peptide, then 10ul ATP/Mg₂₊ (5mM ATP/50mM MgCl₂), then 10ul of 5x Assay Buffer (40mM imidazole hydrochloride, pH7.3, 40 mM beta-glycerophosphate, 1mM EGTA, 100mM MgCl₂, 5 mM MnCl₂, 0.5 mg/ml BSA), then 5ul of Sodium Vanadate(1mM), and then 5ul of water. Mix the components gently and preincubate the reaction mix at 30°C for 2 min. Initial the reaction by adding 10ul of the control enzyme or the filtered supernatant.

The tyrosine kinase assay reaction is then terminated by adding 10 ul of 120mm EDTA and place the reactions on ice.

Tyrosine kinase activity is determined by transferring 50 ul aliquot of reaction mixture to a microtiter plate (MTP) module and incubating at 37°C for 20 min. This allows the streptavadin coated 96 well plate to associate with the biotinylated peptide. Wash the MTP module with 300ul/well of PBS four times. Next add 75 ul of anti-phospotyrosine antibody conjugated to horse radish peroxidase(anti-P-Tyr-POD(0.5u/ml)) to each well and incubate at 37°C for one hour. Wash the well as above.

Next add 100ul of peroxidase substrate solution (Boehringer Mannheim) and incubate at room temperature for at least 5 mins (up to 30 min). Measure the absorbance of the sample at 405 nm by using ELISA reader. The level of bound peroxidase activity is quantitated using an ELISA reader and reflects the level of tyrosine kinase activity.

Example 20: High-Throughput Screening Assay Identifying Phosphorylation Activity

As a potential alternative and/or compliment to the assay of protein tyrosine kinase activity described in Example 19, an assay which detects activation (phosphorylation) of major intracellular signal transduction intermediates can also be used. For example, as described below one particular assay can detect tyrosine

WO 99/47540 PCT/US99/05804

269

phosphorylation of the Erk-1 and Erk-2 kinases. However, phosphorylation of other molecules, such as Raf, JNK, p38 MAP, Map kinase kinase (MEK), MEK kinase, Src, Muscle specific kinase (MuSK), IRAK, Tec, and Janus, as well as any other phosphoserine, phosphotyrosine, or phosphothreonine molecule, can be detected by substituting these molecules for Erk-1 or Erk-2 in the following assay.

Specifically, assay plates are made by coating the wells of a 96-well ELISA plate with 0.1ml of protein G (1ug/ml) for 2 hr at room temp, (RT). The plates are then rinsed with PBS and blocked with 3% BSA/PBS for 1 hr at RT. The protein G plates are then treated with 2 commercial monoclonal antibodies (100ng/well) against

10 Erk-1

15

20

25

5

and Erk-2 (1 hr at RT) (Santa Cruz Biotechnology). (To detect other molecules, this step can easily be modified by substituting a monoclonal antibody detecting any of the above described molecules.) After 3-5 rinses with PBS, the plates are stored at 4°C until use.

A431 cells are seeded at 20,000/well in a 96-well Loprodyne filterplate and cultured overnight in growth medium. The cells are then starved for 48 hr in basal medium (DMEM) and then treated with EGF (6ng/well) or 50 ul of the supernatants obtained in Example 11 for 5-20 minutes. The cells are then solubilized and extracts filtered directly into the assay plate.

After incubation with the extract for 1 hr at RT, the wells are again rinsed. As a positive control, a commercial preparation of MAP kinase (10ng/well) is used in place

of A431 extract. Plates are then treated with a commercial polyclonal (rabbit) antibody (lug/ml) which specifically recognizes the phosphorylated epitope of the Erk-1 and Erk-2 kinases (1 hr at RT). This antibody is biotinylated by standard procedures. The bound polyclonal antibody is then quantitated by successive incubations with Europium-streptavidin and Europium fluorescence enhancing reagent in the Wallac DELFIA instrument (time-resolved fluorescence). An increased fluorescent signal over background indicates a phosphorylation.

15

20

25

30

Example 21: Method of Determining Alterations in a Gene Corresponding to a Polynucleotide

RNA isolated from entire families or individual patients presenting with a phenotype of interest (such as a disease) is be isolated. cDNA is then generated from these RNA samples using protocols known in the art. (See, Sambrook.) The cDNA is then used as a template for PCR, employing primers surrounding regions of interest in SEQ ID NO:X. Suggested PCR conditions consist of 35 cycles at 95°C for 30 seconds; 60-120 seconds at 52-58°C; and 60-120 seconds at 70°C, using buffer solutions described in Sidransky, D., et al., Science 252:706 (1991).

PCR products are then sequenced using primers labeled at their 5' end with T4 polynucleotide kinase, employing SequiTherm Polymerase. (Epicentre Technologies). The intron-exon borders of selected exons is also determined and genomic PCR products analyzed to confirm the results. PCR products harboring suspected mutations is then cloned and sequenced to validate the results of the direct sequencing.

PCR products is cloned into T-tailed vectors as described in Holton, T.A. and Graham, M.W., Nucleic Acids Research, 19:1156 (1991) and sequenced with T7 polymerase (United States Biochemical). Affected individuals are identified by mutations not present in unaffected individuals.

Genomic rearrangements are also observed as a method of determining alterations in a gene corresponding to a polynucleotide. Genomic clones isolated according to Example 2 are nick-translated with digoxigenindeoxy-uridine 5'-triphosphate (Boehringer Manheim), and FISH performed as described in Johnson, Cg. et al., Methods Cell Biol. 35:73-99 (1991). Hybridization with the labeled probe is carried out using a vast excess of human cot-1 DNA for specific hybridization to the corresponding genomic locus.

Chromosomes are counterstained with 4,6-diamino-2-phenylidole and propidium iodide, producing a combination of C- and R-bands. Aligned images for precise mapping are obtained using a triple-band filter set (Chroma Technology, Brattleboro, VT) in combination with a cooled charge-coupled device camera (Photometrics, Tucson, AZ) and variable excitation wavelength filters. (Johnson, Cv.

et al., Genet. Anal. Tech. Appl., 8:75 (1991).) Image collection, analysis and chromosomal fractional length measurements are performed using the ISee Graphical Program System. (Inovision Corporation, Durham, NC.) Chromosome alterations of the genomic region hybridized by the probe are identified as insertions, deletions, and translocations. These alterations are used as a diagnostic marker for an associated disease.

Example 22: Method of Detecting Abnormal Levels of a Polypeptide in a Biological Sample

5

10

15

20

25

A polypeptide of the present invention can be detected in a biological sample, and if an increased or decreased level of the polypeptide is detected, this polypeptide is a marker for a particular phenotype. Methods of detection are numerous, and thus, it is understood that one skilled in the art can modify the following assay to fit their particular needs.

For example, antibody-sandwich ELISAs are used to detect polypeptides in a sample, preferably a biological sample. Wells of a microtiter plate are coated with specific antibodies, at a final concentration of 0.2 to 10 ug/ml. The antibodies are either monoclonal or polyclonal and are produced by the method described in Example 10. The wells are blocked so that non-specific binding of the polypeptide to the well is reduced.

The coated wells are then incubated for > 2 hours at RT with a sample containing the polypeptide. Preferably, serial dilutions of the sample should be used to validate results. The plates are then washed three times with deionized or distilled water to remove unbounded polypeptide.

Next, 50 ul of specific antibody-alkaline phosphatase conjugate, at a concentration of 25-400 ng, is added and incubated for 2 hours at room temperature. The plates are again washed three times with deionized or distilled water to remove unbounded conjugate.

Add 75 ul of 4-methylumbelliferyl phosphate (MUP) or p-nitrophenyl

30 phosphate (NPP) substrate solution to each well and incubate 1 hour at room
temperature. Measure the reaction by a microtiter plate reader. Prepare a standard

WO 99/47540

272

PCT/US99/05804

curve, using serial dilutions of a control sample, and plot polypeptide concentration on the X-axis (log scale) and fluorescence or absorbance of the Y-axis (linear scale). Interpolate the concentration of the polypeptide in the sample using the standard curve.

5

10

15

20

25

30

Example 23: Formulating a Polypeptide

The secreted polypeptide composition will be formulated and dosed in a fashion consistent with good medical practice, taking into account the clinical condition of the individual patient (especially the side effects of treatment with the secreted polypeptide alone), the site of delivery, the method of administration, the scheduling of administration, and other factors known to practitioners. The "effective amount" for purposes herein is thus determined by such considerations.

As a general proposition, the total pharmaceutically effective amount of secreted polypeptide administered parenterally per dose will be in the range of about 1 µg/kg/day to 10 mg/kg/day of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 mg/kg/day, and most preferably for humans between about 0.01 and 1 mg/kg/day for the hormone. If given continuously, the secreted polypeptide is typically administered at a dose rate of about 1 µg/kg/hour to about 50 µg/kg/hour, either by 1-4 injections per day or by continuous subcutaneous infusions, for example, using a mini-pump. An intravenous bag solution may also be employed. The length of treatment needed to observe changes and the interval following treatment for responses to occur appears to vary depending on the desired effect.

Pharmaceutical compositions containing the secreted protein of the invention are administered orally, rectally, parenterally, intracistemally, intravaginally, intraperitoneally, topically (as by powders, ointments, gels, drops or transdermal patch), bucally, or as an oral or nasal spray. "Pharmaceutically acceptable carrier" refers to a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type. The term "parenteral" as used herein refers to modes of administration which include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

25

30

The secreted polypeptide is also suitably administered by sustained-release systems. Suitable examples of sustained-release compositions include semipermeable polymer matrices in the form of shaped articles, e.g., films, or mirocapsules. Sustained-release matrices include polylactides (U.S. Pat. No. 3,773,919, EP 58,481), copolymers of L-glutamic acid and gamma-ethyl-L-glutamate 5 (Sidman, U. et al., Biopolymers 22:547-556 (1983)), poly (2-hydroxyethyl methacrylate) (R. Langer et al., J. Biomed. Mater. Res. 15:167-277 (1981), and R. Langer, Chem. Tech. 12:98-105 (1982)), ethylene vinyl acetate (R. Langer et al.) or poly-D- (-)-3-hydroxybutyric acid (EP 133,988). Sustained-release compositions 10 also include liposomally entrapped polypeptides. Liposomes containing the secreted polypeptide are prepared by methods known per se: DE 3,218,121; Epstein et al., Proc. Natl. Acad. Sci. USA 82:3688-3692 (1985); Hwang et al., Proc. Natl. Acad. Sci. USA 77:4030-4034 (1980); EP 52,322; EP 36,676; EP 88,046; EP 143,949; EP 142,641; Japanese Pat. Appl. 83-118008; U.S. Pat. Nos. 4,485,045 and 4,544,545; and 15 EP 102,324. Ordinarily, the liposomes are of the small (about 200-800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol. percent cholesterol, the selected proportion being adjusted for the optimal secreted polypeptide therapy.

For parenteral administration, in one embodiment, the secreted polypeptide is formulated generally by mixing it at the desired degree of purity, in a unit dosage injectable form (solution, suspension, or emulsion), with a pharmaceutically acceptable carrier, i.e., one that is non-toxic to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation. For example, the formulation preferably does not include oxidizing agents and other compounds that are known to be deleterious to polypeptides.

Generally, the formulations are prepared by contacting the polypeptide uniformly and intimately with liquid carriers or finely divided solid carriers or both. Then, if necessary, the product is shaped into the desired formulation. Preferably the carrier is a parenteral carrier, more preferably a solution that is isotonic with the blood of the recipient. Examples of such carrier vehicles include water, saline, Ringer's

10

15

20

25

30

solution, and dextrose solution. Non-aqueous vehicles such as fixed oils and ethyl oleate are also useful herein, as well as liposomes.

The carrier suitably contains minor amounts of additives such as substances that enhance isotonicity and chemical stability. Such materials are non-toxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, succinate, acetic acid, and other organic acids or their salts; antioxidants such as ascorbic acid; low molecular weight (less than about ten residues) polypeptides, e.g., polyarginine or tripeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids, such as glycine, glutamic acid, aspartic acid, or arginine; monosaccharides, disaccharides, and other carbohydrates including cellulose or its derivatives, glucose, manose, or dextrins; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; counterions such as sodium; and/or nonionic surfactants such as polysorbates, poloxamers, or PEG.

The secreted polypeptide is typically formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml, preferably 1-10 mg/ml, at a pH of about 3 to 8. It will be understood that the use of certain of the foregoing excipients, carriers, or stabilizers will result in the formation of polypeptide salts.

Any polypeptide to be used for therapeutic administration can be sterile. Sterility is readily accomplished by filtration through sterile filtration membranes (e.g., 0.2 micron membranes). Therapeutic polypeptide compositions generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

Polypeptides ordinarily will be stored in unit or multi-dose containers, for example, sealed ampoules or vials, as an aqueous solution or as a lyophilized formulation for reconstitution. As an example of a lyophilized formulation, 10-ml vials are filled with 5 ml of sterile-filtered 1% (w/v) aqueous polypeptide solution, and the resulting mixture is lyophilized. The infusion solution is prepared by reconstituting the lyophilized polypeptide using bacteriostatic Water-for-Injection.

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical

WO 99/47540 PCT/US99/05804

compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration. In addition, the polypeptides of the present invention may be employed in conjunction with other therapeutic compounds.

Example 24: Method of Treating Decreased Levels of the Polypeptide

It will be appreciated that conditions caused by a decrease in the standard or normal expression level of a secreted protein in an individual can be treated by administering the polypeptide of the present invention, preferably in the secreted form. Thus, the invention also provides a method of treatment of an individual in need of an increased level of the polypeptide comprising administering to such an individual a pharmaceutical composition comprising an amount of the polypeptide to increase the activity level of the polypeptide in such an individual.

For example, a patient with decreased levels of a polypeptide receives a daily dose 0.1-100 ug/kg of the polypeptide for six consecutive days. Preferably, the polypeptide is in the secreted form. The exact details of the dosing scheme, based on administration and formulation, are provided in Example 23.

20

25

30

5

10

15

Example 25: Method of Treating Increased Levels of the Polypeptide

Antisense technology is used to inhibit production of a polypeptide of the present invention. This technology is one example of a method of decreasing levels of a polypeptide, preferably a secreted form, due to a variety of etiologies, such as cancer.

For example, a patient diagnosed with abnormally increased levels of a polypeptide is administered intravenously antisense polynucleotides at 0.5, 1.0, 1.5, 2.0 and 3.0 mg/kg day for 21 days. This treatment is repeated after a 7-day rest period if the treatment was well tolerated. The formulation of the antisense polynucleotide is provided in Example 23.

15

20

25

30

WO 99/47540 PCT/US99/05804

276

Example 26: Method of Treatment Using Gene Therapy

One method of gene therapy transplants fibroblasts, which are capable of expressing a polypeptide, onto a patient. Generally, fibroblasts are obtained from a subject by skin biopsy. The resulting tissue is placed in tissue-culture medium and separated into small pieces. Small chunks of the tissue are placed on a wet surface of a tissue culture flask, approximately ten pieces are placed in each flask. The flask is turned upside down, closed tight and left at room temperature over night. After 24 hours at room temperature, the flask is inverted and the chunks of tissue remain fixed to the bottom of the flask and fresh media (e.g., Ham's F12 media, with 10% FBS, penicillin and streptomycin) is added. The flasks are then incubated at 37°C for approximately one week.

At this time, fresh media is added and subsequently changed every several days. After an additional two weeks in culture, a monolayer of fibroblasts emerge. The monolayer is trypsinized and scaled into larger flasks.

pMV-7 (Kirschmeier, P.T. et al., DNA, 7:219-25 (1988)), flanked by the long terminal repeats of the Moloney murine sarcoma virus, is digested with EcoRI and HindIII and subsequently treated with calf intestinal phosphatase. The linear vector is fractionated on agarose gel and purified, using glass beads.

The cDNA encoding a polypeptide of the present invention can be amplified using PCR primers which correspond to the 5' and 3' end sequences respectively as set forth in Example 1. Preferably, the 5' primer contains an EcoRI site and the 3' primer includes a HindIII site. Equal quantities of the Moloney murine sarcoma virus linear backbone and the amplified EcoRI and HindIII fragment are added together, in the presence of T4 DNA ligase. The resulting mixture is maintained under conditions appropriate for ligation of the two fragments. The ligation mixture is then used to transform bacteria HB101, which are then plated onto agar containing kanamycin for the purpose of confirming that the vector has the gene of interest properly inserted.

The amphotropic pA317 or GP+am12 packaging cells are grown in tissue culture to confluent density in Dulbecco's Modified Eagles Medium (DMEM) with 10% calf serum (CS), penicillin and streptomycin. The MSV vector containing the gene is then added to the media and the packaging cells transduced with the vector.

WO 99/47540

5

10

15

20

25

30

The packaging cells now produce infectious viral particles containing the gene (the packaging cells are now referred to as producer cells).

Fresh media is added to the transduced producer cells, and subsequently, the media is harvested from a 10 cm plate of confluent producer cells. The spent media, containing the infectious viral particles, is filtered through a millipore filter to remove detached producer cells and this media is then used to infect fibroblast cells. Media is removed from a sub-confluent plate of fibroblasts and quickly replaced with the media from the producer cells. This media is removed and replaced with fresh media. If the titer of virus is high, then virtually all fibroblasts will be infected and no selection is required. If the titer is very low, then it is necessary to use a retroviral vector that has a selectable marker, such as neo or his. Once the fibroblasts have been efficiently infected, the fibroblasts are analyzed to determine whether protein is produced.

The engineered fibroblasts are then transplanted onto the host, either alone or after having been grown to confluence on cytodex 3 microcarrier beads.

Example 27: Method of Treatment Using Gene Therapy - In Vivo

Another aspect of the present invention is using *in vivo* gene therapy methods to treat disorders, diseases and conditions. The gene therapy method relates to the introduction of naked nucleic acid (DNA, RNA, and antisense DNA or RNA) sequences into an animal to increase or decrease the expression of the polypeptide. The polynucleotide of the present invention may be operatively linked to a promoter or any other genetic elements necessary for the expression of the polypeptide by the target tissue. Such gene therapy and delivery techniques and methods are known in the art, see, for example, WO90/11092, WO98/11779; U.S. Patent NO. 5693622, 5705151, 5580859; Tabata H. et al. (1997) Cardiovasc. Res. 35(3):470-479, Chao J et al. (1997) Pharmacol. Res. 35(6):517-522, Wolff J.A. (1997) Neuromuscul. Disord. 7(5):314-318, Schwartz B. et al. (1996) Gene Ther. 3(5):405-411, Tsurumi Y. et al. (1996) Circulation 94(12):3281-3290 (incorporated herein by reference).

The polynucleotide constructs may be delivered by any method that delivers injectable materials to the cells of an animal, such as, injection into the interstitial space of tissues (heart, muscle, skin, lung, liver, intestine and the like). The polynucleotide constructs can be delivered in a pharmaceutically acceptable liquid or aqueous carrier.

5

10

15

20

25

30

The term "naked" polynucleotide, DNA or RNA, refers to sequences that are free from any delivery vehicle that acts to assist, promote, or facilitate entry into the cell, including viral sequences, viral particles, liposome formulations, lipofectin or precipitating agents and the like. However, the polynucleotides of the present invention may also be delivered in liposome formulations (such as those taught in Felgner P.L. et al. (1995) Ann. NY Acad. Sci. 772:126-139 and Abdallah B. et al. (1995) Biol. Cell 85(1):1-7) which can be prepared by methods well known to those skilled in the art.

The polynucleotide vector constructs used in the gene therapy method are preferably constructs that will not integrate into the host genome nor will they contain sequences that allow for replication. Any strong promoter known to those skilled in the art can be used for driving the expression of DNA. Unlike other gene therapies techniques, one major advantage of introducing naked nucleic acid sequences into target cells is the transitory nature of the polynucleotide synthesis in the cells. Studies have shown that non-replicating DNA sequences can be introduced into cells to provide production of the desired polypeptide for periods of up to six months.

The polynucleotide construct can be delivered to the interstitial space of tissues within the an animal, including of muscle, skin, brain, lung, liver, spleen, bone marrow, thymus, heart, lymph, blood, bone, cartilage, pancreas, kidney, gall bladder, stomach, intestine, testis, ovary, uterus, rectum, nervous system, eye, gland, and connective tissue. Interstitial space of the tissues comprises the intercellular fluid, mucopolysaccharide matrix among the reticular fibers of organ tissues, elastic fibers in the walls of vessels or chambers, collagen fibers of fibrous tissues, or that same matrix within connective tissue ensheathing muscle cells or in the lacunae of bone. It is similarly the space occupied by the plasma of the circulation and the lymph fluid of the lymphatic channels. Delivery to the interstitial space of muscle tissue is preferred for the reasons discussed below. They may be conveniently delivered by injection

WO 99/47540 - PCT/US99/05804

5

10

15

20

25

30

279

into the tissues comprising these cells. They are preferably delivered to and expressed in persistent, non-dividing cells which are differentiated, although delivery and expression may be achieved in non-differentiated or less completely differentiated cells, such as, for example, stem cells of blood or skin fibroblasts. *In vivo* muscle cells are particularly competent in their ability to take up and express polynucleotides.

For the naked polynucleotide injection, an effective dosage amount of DNA or RNA will be in the range of from about 0.05 g/kg body weight to about 50 mg/kg body weight. Preferably the dosage will be from about 0.005 mg/kg to about 20 mg/kg and more preferably from about 0.05 mg/kg to about 5 mg/kg. Of course, as the artisan of ordinary skill will appreciate, this dosage will vary according to the tissue site of injection. The appropriate and effective dosage of nucleic acid sequence can readily be determined by those of ordinary skill in the art and may depend on the condition being treated and the route of administration. The preferred route of administration is by the parenteral route of injection into the interstitial space of tissues. However, other parenteral routes may also be used, such as, inhalation of an aerosol formulation particularly for delivery to lungs or bronchial tissues, throat or mucous membranes of the nose. In addition, naked polynucleotide constructs can be delivered to arteries during angioplasty by the catheter used in the procedure.

The dose response effects of injected polynucleotide in muscle *in vivo* is determined as follows. Suitable template DNA for production of mRNA coding for polypeptide of the present invention is prepared in accordance with a standard recombinant DNA methodology. The template DNA, which may be either circular or linear, is either used as naked DNA or complexed with liposomes. The quadriceps muscles of mice are then injected with various amounts of the template DNA.

Five to six week old female and male Balb/C mice are anesthetized by intraperitoneal injection with 0.3 ml of 2.5% Avertin. A 1.5 cm incision is made on the anterior thigh, and the quadriceps muscle is directly visualized. The template DNA is injected in 0.1 ml of carrier in a 1 cc syringe through a 27 gauge needle over one minute, approximately 0.5 cm from the distal insertion site of the muscle into the knee and about 0.2 cm deep. A suture is placed over the injection site for future localization, and the skin is closed with stainless steel clips.

After an appropriate incubation time (e.g., 7 days) muscle extracts are prepared by excising the entire quadriceps. Every fifth 15 um cross-section of the individual quadriceps muscles is histochemically stained for protein expression. A time course for protein expression may be done in a similar fashion except that quadriceps from different mice are harvested at different times. Persistence of DNA in muscle following injection may be determined by Southern blot analysis after preparing total cellular DNA and HIRT supernatants from injected and control mice. The results of the above experimentation in mice can be use to extrapolate proper dosages and other treatment parameters in humans and other animals using naked DNA.

Example 28: Transgenic Animals.

5

10

15

20

25

30

The polypeptides of the invention can also be expressed in transgenic animals. Animals of any species, including, but not limited to, mice, rats, rabbits, hamsters, guinea pigs, pigs, micro-pigs, goats, sheep, cows and non-human primates, e.g., baboons, monkeys, and chimpanzees may be used to generate transgenic animals. In a specific embodiment, techniques described herein or otherwise known in the art, are used to express polypeptides of the invention in humans, as part of a gene therapy protocol.

Any technique known in the art may be used to introduce the transgene (i.e., polynucleotides of the invention) into animals to produce the founder lines of transgenic animals. Such techniques include, but are not limited to, pronuclear microinjection (Paterson et al., Appl. Microbiol. Biotechnol. 40:691-698 (1994); Carver et al., Biotechnology (NY) 11:1263-1270 (1993); Wright et al., Biotechnology (NY) 9:830-834 (1991); and Hoppe et al., U.S. Pat. No. 4,873,191 (1989)); retrovirus mediated gene transfer into germ lines (Van der Putten et al., Proc. Natl. Acad. Sci., USA 82:6148-6152 (1985)), blastocysts or embryos; gene targeting in embryonic stem cells (Thompson et al., Cell 56:313-321 (1989)); electroporation of cells or embryos (Lo, 1983, Mol Cell. Biol. 3:1803-1814 (1983)); introduction of the polynucleotides of the invention using a gene gun (see, e.g., Ulmer et al., Science 259:1745 (1993); introducing nucleic acid constructs into embryonic pleuripotent stem cells and transferring the stem cells back into the blastocyst; and sperm-

mediated gene transfer (Lavitrano et al., Cell 57:717-723 (1989); etc. For a review of such techniques, see Gordon, "Transgenic Animals," Intl. Rev. Cytol. 115:171-229 (1989), which is incorporated by reference herein in its entirety.

Any technique known in the art may be used to produce transgenic clones containing polynucleotides of the invention, for example, nuclear transfer into enucleated oocytes of nuclei from cultured embryonic, fetal, or adult cells induced to quiescence (Campell et al., Nature 380:64-66 (1996); Wilmut et al., Nature 385:810-813 (1997)).

5

10

15

20

25

30

The present invention provides for transgenic animals that carry the transgene in all their cells, as well as animals which carry the transgene in some, but not all their cells, i.e., mosaic animals or chimeric. The transgene may be integrated as a single transgene or as multiple copies such as in concatamers, e.g., head-to-head tandems or head-to-tail tandems. The transgene may also be selectively introduced into and activated in a particular cell type by following, for example, the teaching of Lasko et al. (Lasko et al., Proc. Natl. Acad. Sci. USA 89:6232-6236 (1992)). The regulatory sequences required for such a cell-type specific activation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art. When it is desired that the polynucleotide transgene be integrated into the chromosomal site of the endogenous gene, gene targeting is preferred. Briefly, when such a technique is to be utilized, vectors containing some nucleotide sequences homologous to the endogenous gene are designed for the purpose of integrating, via homologous recombination with chromosomal sequences, into and disrupting the function of the nucleotide sequence of the endogenous gene. The transgene may also be selectively introduced into a particular cell type, thus inactivating the endogenous gene in only that cell type, by following, for example, the teaching of Gu et al., Science 265:103-106 (1994)). The regulatory sequences required for such a cell-type specific inactivation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art.

Once transgenic animals have been generated, the expression of the recombinant gene may be assayed utilizing standard techniques. Initial screening may be accomplished by Southern blot analysis or PCR techniques to analyze animal tissues to verify that integration of the transgene has taken place. The level of mRNA

10

15

20

30

WO 99/47540 PCT/US99/05804

282

expression of the transgene in the tissues of the transgenic animals may also be assessed using techniques which include, but are not limited to, Northern blot analysis of tissue samples obtained from the animal, in situ hybridization analysis, and reverse transcriptase-PCR (rt-PCR). Samples of transgenic gene-expressing tissue may also be evaluated immunocytochemically or immunohistochemically using antibodies specific for the transgene product.

Once the founder animals are produced, they may be bred, inbred, outbred, or crossbred to produce colonies of the particular animal. Examples of such breeding strategies include, but are not limited to: outbreeding of founder animals with more than one integration site in order to establish separate lines; inbreeding of separate lines in order to produce compound transgenics that express the transgene at higher levels because of the effects of additive expression of each transgene; crossing of heterozygous transgenic animals to produce animals homozygous for a given integration site in order to both augment expression and eliminate the need for screening of animals by DNA analysis; crossing of separate homozygous lines to produce compound heterozygous or homozygous lines; and breeding to place the transgene on a distinct background that is appropriate for an experimental model of interest.

Transgenic animals of the invention have uses which include, but are not limited to, animal model systems useful in elaborating the biological function of polypeptides of the present invention, studying conditions and/or disorders associated with aberrant expression, and in screening for compounds effective in ameliorating such conditions and/or disorders.

25 Example 29: Knock-Out Animals.

Endogenous gene expression can also be reduced by inactivating or "knocking out" the gene and/or its promoter using targeted homologous recombination. (*E.g.*, see Smithies et al., Nature 317:230-234 (1985); Thomas & Capecchi, Cell 51:503-512 (1987); Thompson et al., Cell 5:313-321 (1989); each of which is incorporated by reference herein in its entirety). For example, a mutant, non-functional polynucleotide of the invention (or a completely unrelated DNA sequence) flanked by DNA homologous to the endogenous polynucleotide sequence (either the coding

15

20

25

30

regions or regulatory regions of the gene) can be used, with or without a selectable marker and/or a negative selectable marker, to transfect cells that express polypeptides of the invention in vivo. In another embodiment, techniques known in the art are used to generate knockouts in cells that contain, but do not express the gene of interest. Insertion of the DNA construct, via targeted homologous recombination, results in inactivation of the targeted gene. Such approaches are particularly suited in research and agricultural fields where modifications to embryonic stem cells can be used to generate animal offspring with an inactive targeted gene (e.g., see Thomas & Capecchi 1987 and Thompson 1989, supra). However this approach can be routinely adapted for use in humans provided the recombinant DNA constructs are directly administered or targeted to the required site in vivo using appropriate viral vectors that will be apparent to those of skill in the art.

In further embodiments of the invention, cells that are genetically engineered to express the polypeptides of the invention, or alternatively, that are genetically engineered not to express the polypeptides of the invention (e.g., knockouts) are administered to a patient in vivo. Such cells may be obtained from the patient (i.e., animal, including human) or an MHC compatible donor and can include, but are not limited to fibroblasts, bone marrow cells, blood cells (e.g., lymphocytes), adipocytes, muscle cells, endothelial cells etc. The cells are genetically engineered in vitro using recombinant DNA techniques to introduce the coding sequence of polypeptides of the invention into the cells, or alternatively, to disrupt the coding sequence and/or endogenous regulatory sequence associated with the polypeptides of the invention, e.g., by transduction (using viral vectors, and preferably vectors that integrate the transgene into the cell genome) or transfection procedures, including, but not limited to, the use of plasmids, cosmids, YACs, naked DNA, electroporation, liposomes, etc. The coding sequence of the polypeptides of the invention can be placed under the control of a strong constitutive or inducible promoter or promoter/enhancer to achieve expression, and preferably secretion, of the polypeptides of the invention. The engineered cells which express and preferably secrete the polypeptides of the invention can be introduced into the patient systemically, e.g., in the circulation, or intraperitoneally.

WO 99/47540 PCT/US99/05804

284

Alternatively, the cells can be incorporated into a matrix and implanted in the body, e.g., genetically engineered fibroblasts can be implanted as part of a skin graft; genetically engineered endothelial cells can be implanted as part of a lymphatic or vascular graft. (See, for example, Anderson et al. U.S. Patent No. 5,399,349; and Mulligan & Wilson, U.S. Patent No. 5,460,959 each of which is incorporated by reference herein in its entirety).

When the cells to be administered are non-autologous or non-MHC compatible cells, they can be administered using well known techniques which prevent the development of a host immune response against the introduced cells. For example, the cells may be introduced in an encapsulated form which, while allowing for an exchange of components with the immediate extracellular environment, does not allow the introduced cells to be recognized by the host immune system.

10

15

20

25

Transgenic and "knock-out" animals of the invention have uses which include, but are not limited to, animal model systems useful in elaborating the biological function of polypeptides of the present invention, studying conditions and/or disorders associated with aberrant expression, and in screening for compounds effective in ameliorating such conditions and/or disorders.

It will be clear that the invention may be practiced otherwise than as particularly described in the foregoing description and examples. Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, are within the scope of the appended claims.

The entire disclosure of each document cited (including patents, patent applications, journal articles, abstracts, laboratory manuals, books, or other disclosures) in the Background of the Invention, Detailed Description, and Examples is hereby incorporated herein by reference. Further, the hard copy of the sequence listing submitted herewith and the corresponding computer readable form are both incorporated herein by reference in their entireties.

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

on page 180 , line B. IDENTIFICATIONOFDEPOSIT Name of depositary institution American Type Culture Collect Address of depositary institution (including postal code and country)	Further deposits are identified on an additional sheet			
Name of depositary institution American Type Culture Collect				
	stion			
Address of depositary institution (including postal code and country				
10801 University Boulevard Manassas, Virginia 20110-2209 United States of America	y)			
Date of deposit	Accession Number			
February 12, 1998	209628			
C. ADDITIONAL INDICATIONS (leave blank if not applicable)	This information is continued on an additional sheet			
D. DESIGNATED STATES FOR WHICH INDICATIONS	S ARE MADE (if the indications are not for all designated States)			
E. SEPARATE FURNISHING OF INDICATIONS (leave bla	ank if not applicable)			
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")				
For receiving Office use only	For International Bureau use only			
This sheet was received with the international application Authorized officer Form PCT/RO/134 (July 1992)	This sheet was received by the International Bureau on: Authorized officer			

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorgan		
on page, me		
B. IDENTIFICATIONOFDEPOSIT	Further deposits are identified on an additional sheet	
Name of depositary institution American Type Cultt	ture Collection	
Address of depositary institution (including postal code	de and country)	
10801 University Boulevard Manassas, Virginia 20110-2209		
United States of America		
Date of deposit	Accession Number	
February 25, 1998	209641	
C. ADDITIONAL INDICATIONS (leave blank if no	not applicable) This information is continued on an additional sheet	
-		
·		
D. DESIGNATED STATES FOR WHICH IND	DICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIO	ONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")		
For receiving Office use only	For International Bureau use only	
This sheet was received with the international appli	lication This sheet was received by the International Bureau on:	
Auborged officer	Authorized officer	
Authorized officer	Authorized officer	

Form PCT/RO/134 (July 1992)

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism reference on page186, line	red to in the description N/A .
B. IDENTIFICATIONOFDEPOSIT	Further deposits are identified on an additional sheet
Name of depositary institution American Type Culture Colle	ction
Address of depositary institution (including postal code and count 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America	(ry)
Date of deposit	Accession Number
March 4, 1998	209651
C. ADDITIONAL INDICATIONS (leave blank if not applicable)	This information is continued on an additional sheet
D. DESIGNATED STATES FOR WHICH INDICATION	
E. SEPARATE FURNISHING OF INDICATIONS (leave The indications listed below will be submitted to the Internation Number of Deposit")	nal Bureau later (specify the general nature of the indications e.g., "Accession
	For International Pursuance only
For receiving Office use only This sheet was received with the international application () () Authorized officer	For International Bureau use only This sheet was received by the International Bureau on: Authorized officer

Form PCT/RO/134 (July 1992)

10

15

20

What Is Claimed Is:

- 1. An isolated nucleic acid molecule comprising a polynucleotide having a nucleotide sequence at least 95% identical to a sequence selected from the group consisting of:
- (a) a polynucleotide fragment of SEQ ID NO:X or a polynucleotide fragment of the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
- (b) a polynucleotide encoding a polypeptide fragment of SEQ ID NO:Y or a polypeptide fragment encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
 - (c) a polynucleotide encoding a polypeptide domain of SEQ ID NO:Y or a polypeptide domain encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
- (d) a polynucleotide encoding a polypeptide epitope of SEQ ID NO:Y or a polypeptide epitope encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
- (e) a polynucleotide encoding a polypeptide of SEQ ID NO:Y or the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X, having biological activity;
 - (f) a polynucleotide which is a variant of SEQ ID NO:X;
 - (g) a polynucleotide which is an allelic variant of SEQ ID NO:X;
- (h) a polynucleotide which encodes a species homologue of the SEQ ID NO:Y;
- (i) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h), wherein said polynucleotide does not hybridize under stringent conditions to a nucleic acid molecule having a nucleotide sequence of only A residues or of only T residues.

- 2. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding a secreted protein.
- 5 3. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding the sequence identified as SEQ ID NO:Y or the polypeptide encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X.
- 10 4. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises the entire nucleotide sequence of SEQ ID NO:X or the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X.
- 15 5. The isolated nucleic acid molecule of claim 2, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.
- 6. The isolated nucleic acid molecule of claim 3, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.
 - 7. A recombinant vector comprising the isolated nucleic acid molecule of claim 1.

- 8. A method of making a recombinant host cell comprising the isolated nucleic acid molecule of claim 1.
 - 9. A recombinant host cell produced by the method of claim 8.

30

10. The recombinant host cell of claim 9 comprising vector sequences.

PCT/US99/05804

- 11. An isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence selected from the group consisting of:
- (a) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence5 included in ATCC Deposit No:Z;
 - (b) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z, having biological activity;
 - (c) a polypeptide domain of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;
- (d) a polypeptide epitope of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;
 - (e) a secreted form of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;
- (f) a full length protein of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;
 - (g) a variant of SEQ ID NO:Y;

WO 99/47540

- (h) an allelic variant of SEQ ID NO:Y; or
- (i) a species homologue of the SEQ ID NO:Y.
- The isolated polypeptide of claim 11, wherein the secreted form or the
 full length protein comprises sequential amino acid deletions from either the C-terminus or the N-terminus.
 - 13. An isolated antibody that binds specifically to the isolated polypeptide of claim 11.
 - 14. A recombinant host cell that expresses the isolated polypeptide of claim 11.
 - 15. A method of making an isolated polypeptide comprising:
- 30 (a) culturing the recombinant host cell of claim 14 under conditions such that said polypeptide is expressed; and

- (b) recovering said polypeptide.
- 16. The polypeptide produced by claim 15.
- 5 17. A method for preventing, treating, or ameliorating a medical condition, comprising administering to a mammalian subject a therapeutically effective amount of the polypeptide of claim 11 or the polynucleotide of claim 1.
- 18. A method of diagnosing a pathological condition or a susceptibility to
 10 a pathological condition in a subject comprising:
 - (a) determining the presence or absence of a mutation in the polynucleotide of claim 1; and
 - (b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or absence of said mutation.

- 19. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising:
- (a) determining the presence or amount of expression of the polypeptide of claim 11 in a biological sample; and
- 20 (b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or amount of expression of the polypeptide.
 - 20. A method for identifying a binding partner to the polypeptide of claim 11 comprising:
- 25 (a) contacting the polypeptide of claim 11 with a binding partner; and
 - (b) determining whether the binding partner effects an activity of the polypeptide.
 - 21. The gene corresponding to the cDNA sequence of SEQ ID NO:Y.

- 22. A method of identifying an activity in a biological assay, wherein the method comprises:
 - (a) expressing SEQ ID NO:X in a cell;
 - (b) isolating the supernatant;
- 5 (c) detecting an activity in a biological assay; and
 - (d) identifying the protein in the supernatant having the activity.
 - 23. The product produced by the method of claim 20.

```
<110> Human Genome Sciences, Inc., et al.
```

- <120> 95 Human Secreted Proteins
- <130> PZ027PCT
- <140> Unassigned
- <141> 1999, March 18
- <150> 60/078,566
- <151> 1998-03-19
- <150> 60/078,574
- <151> 1998-03-19
- <150> 60/078,576
- <151> 1998-03-19
- <150> 60/078,563
- <151> 1998-03-19
- <150> 60/078,573
- <151> 1998-03-19
- <150> 60/078,578
- <151> 1998-03-19
- <150> 60/078,579
- <151> 1998-03-19
- <150> 60/078,581
- <151> 1998-03-19
- <150> 60/078,577
- <151> 1998-03-19
- <150> 60/080,314
- <151> 1998-04-01
- <150> 60/080,312
- <151> 1998-04-01
- <150> 60/080,313
- <151> 1998-04-01
- <160> 392
- <170> PatentIn Ver. 2.0
- <210> 1
- <211> 733
- <212> DNA
- <213> Homo sapiens
- <400> 1

PCT/US99/05804

```
gggatccgga gcccaaatct tctgacaaaa ctcacacatg cccaccgtgc ccagcacctg
                                                                         60
aattcgaggg tgcaccgtca gtcttcctct tccccccaaa acccaaggac accctcatga
                                                                        120
                                                                        180
tctcccggac tcctgaggtc acatgcgtgg tggtggacgt aagccacgaa gaccctgagg
                                                                        240
tcaagttcaa ctggtacgtg gacggcgtgg aggtgcataa tgccaagaca aagccgcggg
aggagcagta caacagcacg taccgtgtgg tcagcgtcct caccgtcctg caccaggact
                                                                        300
ggctgaatgg caaggagtac aagtgcaagg tctccaacaa agccctccca acccccatcg
                                                                        360
agaaaaccat ctccaaagcc aaagggcagc cccgagaacc acaggtgtac accctgcccc
                                                                        420
                                                                        480
catcceggga tgagctgacc aagaaccagg tcagcctgac ctgcctggtc aaaggcttct
                                                                        540
atccaagcga catcgccgtg gagtgggaga gcaatgggca gccggagaac aactacaaga
ccacgcctcc cgtgctggac tccgacggct ccttcttcct ctacagcaag ctcaccgtgg
                                                                        600
acaagagcag gtggcagcag gggaacgtct tctcatgctc cgtgatgcat gaggctctgc
                                                                        660
acaaccacta cacgcagaag agcctctccc tgtctccggg taaatgagtg cgacggccgc
                                                                        720
gactctagag gat
                                                                        733
<210> 2
<211> 5
<212> PRT
<213> Homo sapiens
<220>
<221> Site
<222> (3)
<223> Xaa equals any of the twenty naturally ocurring L-amino acids
<400> 2
Trp Ser Xaa Trp Ser
<210> 3
<211> 86
<212> DNA
<213> Homo sapiens
<400> 3
gcgcctcgag atttccccga aatctagatt tccccgaaat gatttccccg aaatgatttc
                                                                         60
                                                                         86
cccgaaatat ctgccatctc aattag
<210> 4
<211> 27
<212> DNA
<213> Homo sapiens
<400> 4
                                                                         27
gcggcaagct ttttgcaaag cctaggc
<210> 5
<211> 271
<212> DNA
<213> Homo sapiens
<400> 5
ctcgagattt ccccgaaatc tagatttccc cgaaatgatt tccccgaaat gatttccccg 60
```

PCT/US99/05804

gcccctaact ttatgcagag	ccgcccagtt gccgaggccg	agtcagcaac ccgcccattc cctcggcctc gcaaaaagct	tccgcccat tgagctattc	ggctgactaa	tttttttat	120 180 240 271
<210> 6 <211> 32 <212> DNA <213> Homo	sapiens					
<400> 6 gcgctcgagg	gatgacagcg	atagaacccc	gg			32
<210> 7 <211> 31 <212> DNA <213> Homo	sapiens					
<400> 7	-	ggatccgcct	С			31
<210> 8 <211> 12 <212> DNA <213> Homo	sapiens					
<400> 8 ggggactttc	cc					12
<210> 9 <211> 73 <212> DNA <213> Homo	sapiens					
<400> 9 gcggcctcga ccatctcaat		ccggggactt	tccggggact	ttccgggact	ttccatcctg	60 73
<210> 10 <211> 256 <212> DNA <213> Homo	sapiens					
caattagtca cagttccgcc	gcaaccatag cattctccgc gcctctgagc	gactttccgg tcccgccct cccatggctg tattccagaa	aactccgccc actaattttt	atcccgcccc tttatttatg	taactccgcc cagaggccga	60 120 180 240 256

.

```
<210> 11
<211> 2343
<212> DNA
<213> Homo sapiens
<400> 11
acgcgtccgg tttttcaaag gtttaactgt ccagggcaga tacttaagac tatctgatca
                                                                       60
tccattaaaa acttttcaca tagtcttgct taaatggatc cattatgttt acccattata
                                                                      120
ttgttctcag ctgtagtttt aagaaattta tttcatttgt taatacttac tttccattac
                                                                      180
cttccccttt tctgtgacaa tccgttgata cttgaagacc tctcttgtat tcatcttagg
                                                                      240
                                                                      300
gttaatattt ttaaggccaa acagcctaaa ttctatggta atcaactcca gccttgtgta
                                                                      360
atgaaatctt ctgcataaag ataggtttaa atcaaatcag attgcagatt ttattgaaga
aattgtgttt ttaagagttg acaaatatat gttgtatggc taaaacaaag aaaatacttc
                                                                      420
tgttgcttct gcatttagta gaagaaaaac tatatatgtt tgtgaccaaa gtataaaata
                                                                      480
tgattctttc cagggaggta aaggttatgc acaagatttt cactagcagc tctaaaaggc
                                                                      540
taccctcaat taattgccat gaacatttca tagccctaga aggatgtagg ctcatttcag
                                                                      600
tgtcatcctg gtttattctt tattgtatta ttcagcagtc attttaacac tatgctagac
                                                                      660
actttagaga ttcagaagag taacagggtt tctgttctca tgaagcttat caggagacag
                                                                      720
aaaacatatg aattagatct aattggaggc aaactgaaat atatagtgga gttagtgtgg
                                                                      780
ttatcagcac ataaatgagt gatccatcaa caaaaggaga aattgggagg gttttatggg
                                                                      840
ccaaaaacag catgattaaa tgtgatagag tatatgtcat gttttaggtg tgatgaacat
                                                                      900
tcagttatgt gtgacgaata ggataattga aaaaatatga aaggctatga tgccagaaag
                                                                      960
tattatggga caagatctta aaaccagtgt tacctaggga gtatgaattt aatatgggaa
                                                                     1020
ttcttaaact cctttatgac tggaagatga gcatcagagt gtctgcgacc attttgatga
                                                                     1080
tatgatgtac cagtttttaa atgtttggct ttttccaggt gatgaaagcg ggggatgagt
                                                                     1140
taagaaccac tgctgtgaag gattcacaac tatttttagg cagttgggta aaaatgacca
                                                                     1200
atttagtttt aagaaactga ctgtggctcc agagtatgtt ggagaagtga aaatggagac
                                                                     1260
taggaataac aggtgggaga ctattagtct aattaagatg taattataaa tctaagctag
                                                                     1320
gaacgtaaaa tgagaatgca aagtaagaaa caaatatggg gaaaattata tgtaaaagta
                                                                     1380
ataggacttg gcatcttact gatgtgattg attatgagaa aaatgaagca tgtggaggag
                                                                     1440
tccactggac agtaggaaat tcagcctaag acttgggtaa gagttctgtg gagttgtgaa
                                                                     1500
ttcagaggcc agagatgtga tatttaaaat tttggttcaa gatttcccag gtataagaaa
                                                                     1560
gcaagaggat taaagcattg taattaaact ttaagcagtg catatttatg ttatagataa
                                                                     1620
gataaacaag aaatctaggg atcaaatagg attaaaatta gtagtgatca ttcagtacag
                                                                     1680
tagttacgta ctgttattca caagagtata taaatcaaat tacaaggaat taaggatata
                                                                     1740
                                                                     1800
aacgtgataa gaaagtatgc actgtactct ttgaggaagt ttgccataga aaggaagaag
                                                                     1860
aaataggatg gtagatcaga agtaaagcag gacccagtgg ggggagtgtt tgcagtgagg
                                                                     1920
cagtatgtat aatcatttaa aacatgggtt tggagtcctc tcaggttcca tgtttgtaat
ggacataatg ataataatcc ctttcattta aggctgttgt gaggattaaa tgtgttaatg
                                                                     1980
tgcaaataac tttacacagt gcctggtata taataaatgc ttgctaccta ttaactagta
                                                                     2040
tttgtttcta aggctaattt aagtcctaga attgattgca aggattagat caggagtata
                                                                     2100
gtggacatgt tgggatttaa atatttaaat atagagatgc tttttaggac cattgttaga
                                                                     2160
accagaagag attttttacc aagttcacac agaaatgtag gtgcattggc tgggcatggt
                                                                     2220
ggctcacacc tgcagtccca gcacttggga aggctgaggc agaagaactg cttgaggcca
                                                                     2280
2340
                                                                     2343
aaa
```

<210> 12

<211> 1177

<212> DNA

<213> Homo sapiens

```
<221> SITE
<222> (1095)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (1115)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (1142)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (1162)
<223> n equals a,t,g, or c
<400> 12
agccaccatg cccggcctag attaaaaatt tgaagacata ttctctacta tgagccaatg
                                                                      60
                                                                     120
aaattactca ttttgtttct atcccatttg ctgtcccttg cttttggaat tttgtgtctt
                                                                     180
agtgtgactg tgattctttc tctccttttg tctttcagca aacggggatt cagcgtccga
tcctttggaa cagggactca cgtgaagctt ccaggaccag ctcccgacaa gcccaatgtt
                                                                     240
                                                                     300
tatgatttca aaaccacata tgaccagatg tacaatgatc ttcttaggaa agacaaagaa
ctctatacac agaatgggat tttacatatg ctggacagaa ataagagaat caagccccgg
                                                                     360
                                                                     420
ccagaaagat tccagaactg caaagacctg tttgatctga tcctcacttg cgaagagaga
                                                                     480
gtgtatgacc aggtggtgga agatctgaat tccagagaac aggagacctg ccagccygtg
                                                                     540
cacgtggtca atgtggacat ccaggacaac cacgaggagg ccaccctggg ggcgtttctc
atctgtgagc tctgccagtg tatccagcac acggaagaca tggagaacga gatcgacgag
                                                                     600
ctgctgcagg agttcgagga gaagagtggc cgcacctttc tgcacaccgt ctgcttctac
                                                                     660
tgagcccagc gcccgcatgg agccgcctct ggagcttcct gttgttcata ctttttcctt
                                                                     720
                                                                     780
cctgacattt gtttttactt acaggtgttc tgctggtgac ggtagcatta cccaaataaa
                                                                     840
ctgtgcatat gaaatgggag aggagatgcc aaaacgccag atgaaagcaa tcaagtttct
                                                                     900
tcttttccac ttttacttat gagcrggata ttgattacaa agtttttctt ctttaaccaa
                                                                     960
aaaggaaaga caacggtttg tgtgcacttc ccgacatacc tgtgtcttcg tgtgcctgcc
                                                                    1020
ttecetecet ectececace gggeeggact gtacagagee etgetgegge gtgttaggaa
                                                                    1080
1140
aaaaaaaaaa raaancaaaa aaaaaaaaaa aaggnggggc cgaaggtttt ttccctttgg
                                                                    1177
tnggggttat ttttggcttg gnattggcct tcgtttt
<210> 13
<211> 2107
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (149)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (487)
```

<223> n equals a,t,g, or c

(223) II eq	uais a,c,g,	or c				
<400> 13						
	catataaaag	aaaacaaaat	atttaaaaca	cttaaaggag	atctgacgaa	60
	aagaaaaata					120
	cctttatcca					180
	caagagactg					240
	ggcgctgatg					300
	ctatgtgaca					360
	tccttttatt					420
	tcaaaaccct					480
	gctggaactg					540
	ggccgtcctt					600
	agttttcact					660
	gacttcattt					720
	accttactca					780
	gtcgcagcga					840
	atgacatctg					900
	ttcattactt					960
	acaygttcta					1020
	tgagagtcca					1080
	tgatgaaagg					1140
	gatgtgggaa					1200
	aactcctgag					1260
	gcaaggcgca					1320
	agacctcagc					1380
gggagattta	gtgacaatgg	tttcctcatg	actgagaaga	cacgaactat	attattttac	1440
ccttgggagc	ccaccagcaa	gaatgagttg	gagcaatctt	ttcatgtgac	ctccttaaca	1500
gatatttact	gaaggaatct	aggttgtatt	ttcagtggac	aatgggaata	aagcatttct	1560
aaagcaccga	ctggagagga	aggcaacaga	gacaaggaga	gaagccgaga	gacatgtctg	1620
cgtgctgcca	cgcatctgag	cgattgctct	gtgaagagtt	gtacactgaa	cattttcagg	1680
ggaggctgtt	tacccaggca	atgtcctcaa	acaagcctgt	gccggggtgt	cctggaatct	1740
gtgccaggac	tgtgttttta	gcccttcacc	tctcagcttt	agcaggacat	gaaccagtta	1800
taacaagatg	sccctgcagc	tggttacaag	aatgtgacat	ggcaggatct	atggaaccaa	1860
atggaaggtt	ttgaggtgat	gtaggtcttt	cacagttagc	tttggggaat	acagaatact	1920
caaataaagt	gctttgttat	tatttcagag	ggaatggcga	ttgaaatgtt	acaacagaga	1980
tttcttggtg	gtagctattt	gggtaaaggt	atatggatat	ttttctgtac	atgtgaaatt	2040
atataaaaat	aaaagttata	taaattacat	tgaaaaaaaa	aaaaaaaaa	aaaaaaggg	2100
cggccgc						2107
<210> 14						
<211> 1262						
<212> DNA						
<213> Homo	sapiens					
<400> 14						
	cgasctgaat					60
	gacaaagagg					120
	aatgattgga					180
	gaatgaagtg					240
	ccccatatcc					300
	tggctgtgct					360
	agacctttgt					420
gcatgcagcc	ctctatcatt	tttggctccg	agaagaggc	aaggggcccc	cgcaggtarc	480

7

```
ttctgtgctt gccctcgccc tgccagcagg cagctgtgcc cctggcctgc ccttcccggg
                                                                    540
accepttatt ccaacteage teetetttge actggaatgg ggcactecaa cacceptcag
                                                                     600
ggaccaccct ccccacagta tgcactcagc cccacagaac ccaccagtct ttctgggaac
                                                                     660
                                                                    720
tcacacctgc ccgccatctt ggtactttag gttaatccct caagcatgaa agctggatct
                                                                    780
tttggggttt aagaagecea ageettgtte etgeeetgge etagggagea eteaggaggg
ttccttggtc ctcatctctc ccacctccgt tccctctggg ccccacacta gccacagcgc
                                                                    840
gggccttgtg ctggagtttg agcctgggac agggagaggg aggcttggag acagtctgac
                                                                    900
                                                                    960
1020
gctctttccc ctttctgggc ctgggtctcc ccatctcttc aatggggctg ataccttcac
                                                                   1080
agcccacage atgggcactt atgaggacaa agtgaattta acctggaaaa gaatgtattt
                                                                   1140
gagagtttct tttaaataat cagcgggtgt tggtgatttg tagcccttct gcccttaaat
                                                                   1200
gcttccttgg gcaagagctg tctgtcctcc ctgcaggagg ctgagtgtga agagtatcat
                                                                   1260
tcattgtttc tctattaaat tattttctgc taaaaaaaaa aaaaaaaaat ttctgcggtc
                                                                   1262
cg
<210> 15
<211> 759
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (16)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (22)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (36)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (51)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (52)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (57)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (58)
```

<223> n equals a,t,g, or c

<400> 15					
ggattaacaa attttncaca	cnaggaaaac	aggttnttga	cccaattagg	nnttttnnca	60
aaaaagctta tttttaggtt	gacacttatt	agaagttacg	ccttgcaggt	taccggttcc	120
ggaattcccg ggtcgaaccc	caaggggttc	gcggacccca	gacatgagga	ggctcctcct	180
ggtcaccagc ctggtggttg	tgctgctgtg	ggaggcaggt	gcagtcccag	cacccaaggt	240
ccctatcaag atgcaagtca	aacactggcc	ctcagagcag	gacccagaga	aggcctgggg	300
cgcccgtgtg gtggagcctc	cggagaagga	cgaccagctg	gtggtgctgt	tccctgtcca	360
gaagccgaaa ctcttgacca	ccgaggagaa	gccacgaggt	cagggcaggg	gccccatcct	420
tccaggcacc aaggcctgga	tggagaccga	ggacaccctg	ggccgtgtcc	tgagtcccga	480
gcccgaccat gacagcctgt					540
ccggttgtgg gtgatgccaa	atcaccaggt	gctcctggga	ccggaggaag	accaagacca	600
catctaccac ccccagtagg					660
caggctgttg ggactgggac	cctccctacc	ctgccccagc	tagacaaata	aaccccagca	720
ggccgggaaa aaaaaaaaa	aaaaaaaag	ggcggccgc			759
.210- 16					
<210> 16					
<211> 1810 <212> DNA					
<213> Homo sapiens					
<400> 16					
cacgagggtg tgcgtgctta	ggcaggaacc	cagttttact	ttatgccatg	tggaaagttt	60
cttttccag tatcaccagt					120
ttacttgcag acttcccacg					180
ctacactgcc tatttccctt					240
gtgttgctga gaatttcctt					300
ggcaggaatt atgccatctg					360
aattagcaga tactcggtaa					420
cacagcactg ccctgggaag					480
accetgecag aactecetgg	atggaagtga	ccaccgtgta	tctgtggatt	gtctgcaggg	540
ctctgctggg gtcagcaggt	cccacaacag	agccagggct	cggtctcctc	atggctgtca	600
gaggtttacg tattccgcct	cctcccacca	aagtctgaag	ttgttgtatt	ccattccttg	660
ctatatccac atcttttaat	aatgctaaaa	tcccgtgttt	ctctaaagca	ttggattgaa	720
ccaactgaag aaggaccacg	tgtgttgctg	ggcctgcttg	ggcacaagcc	gtttccgatc	780
caagtcaact gctggtctgc					840
ggacagcagg tgagaccata					900
agccagagac cctcgtgcac	cagccttcca	ggtgcttatt	ggaacttatg	tcagcccagg	960
ccatatccaa gtgtgtgatg					1020
cctgccctgg tggagctgga					1080
aggatcctga tacaggctgg					1140
gctccaccca gaggggcttg					1200
gatcagtgag cgtccaggag					1260
tgggcctgtg agggcccaca					1320
agcagccgct tttcccagcc					1380
aaagactctc tctaacgtta					1440
agecgcagec tgttcctcac					1500
gatcaccgtg tcttctcgtt					1560
ctttgctctg aagttttgac					1620
cggtgtaacg tgccgttttt					1680
gagttgtctg tggacagtca					1740
tcttctaact gttgcgtatg	gaccaattac	atctgccctt	tctcttccct	gctaaaaaaa	1800
aaaaaaaaa					1810

```
<210> 17
<211> 1052
<212> DNA
<213> Homo sapiens
<400> 17
gcaattttct gcatagcatc agcaatgagt ctgtacaact gtcttgctgc actaattcat
                                                                       60
aagataccat atggacaatg cacgattgca tgtcgtggca aaaacatgga agtgagactt
                                                                      120
atttttctct ctggactgtg catagcagta gctgttgttt gggctgtgtt tcgaaatgaa
                                                                      180
gacaggtggg cttggatttt acaggatatc ttggggattg ctttctgtct gaatttaatt
                                                                      240
                                                                      300
aaaacactga agttgcccaa cttcaagtca tgtgtgatac ttctaggcct tctcctcctc
                                                                      360
tatgatgtat tttttgtttt cataacacca ttcatcacaa agaatggtga gagtatcatg
                                                                     420
gttgaactcg cagctggacc ttttggaaat aatgaaaagt tgccagtagt catcagagta
                                                                      480
ccaaaactga tctatttctc agtaatgagt gtgtgcctca tgcctgtttc aatattgggt
                                                                      540
tttggagaca ttattgtacc aggcctgttg attgcatact gtagaagatt tgatgttcag
                                                                      600
actggttctt cttacatata ctatgtttcg tctacagttg cctatgctat tggcatgata
                                                                      660
cttacatttg ttgttctggt gctgatgaaa aaggggcaac ctgctctcct ctatttagta
                                                                      720
ccttgcacac ttattactgc ctcagttgtt gcctggagac gtaaggaaat gaaaaagttc
                                                                      780
tggaaaggta acagctatca gatgatggac catttggatt gtgcaacaaa tgaagaaaac
                                                                     .840
cctgtgatat ctggtgaaca gattgtccag caataatatt atgtggaact gctataatgt
gtcattgatt ttctacaaat agacttcgac tttttaaatt gacttttgaa ttgacaatct
                                                                      900
                                                                      960
gaaagagtct tcaatgatat gcttgcaaaa atatatttt atgagctggt actgacagtt
                                                                     1020
acatcataaa taactaaaac getttgettt taatgttaaa gttgtgeett cacattaaat
                                                                     1052
aaaacatatg gtctgtgtar tttcaaaaaa aa
<210> 18
<211> 1130
<212> DNA
<213> Homo sapiens
<400> 18
ggcacgaggc catttgtata attctttagt aaattgtatt aatgggagaa tctgtaagtt
                                                                       60
atgtctgaac tttcaggttg tcttataatt gtctttttcc ttatgtcaga tgttctatgt
                                                                      120
cataaqaata aaatggttca caccaataca agtacttagt tgtggaaagg gagagtagaa
                                                                      180
gataaaaatg gagattttcc tgtgctacag gcttagtcaa gcttatggtc tatttaatgg
                                                                      240
ttatcaaaqq caattaaata gtgttgaatg ttctgctttt acctacattt catttttcat
                                                                      300
gtacttagtt acaaattgaa ccctcttcta tttttttcct gctcctgttt ctgtttcatt
                                                                      360
                                                                      420
ttagttttcc ttttccctga ttatcattta ggcatgtaag tgacacccag tagcattgct
ttaattctqc tqqtqacagt gccaaagctt tactatactc tttttgttgt ctgttgcttt
                                                                      480
tctcttgcta atttgcttga ctagataact aagaattcag gtaagcatta gctctttgtt
                                                                      540
cactgagaat aatacaactt gcaagataat taatttggat tgttctacat gtatttcgtt
                                                                      600
tatttctctt taccttgttc atttattacg acattttgaa ttatttacat acccatattt
                                                                      660
                                                                      720
cttctttctt ttatggctca gctcactatg ctttttttta atactggtag cttcctcaag
                                                                      780
gttggaaaac aagatctgaa tactatagaa aataataact atttttctgt ggtcatatta
                                                                      840
aagatataat ggctttggat tttggggtga tttttctact gtcagtttaa aaaaaacttg
                                                                      900
totatttgca tttgtgtgtt attacttcta gttaagagta tttccaagga aagtttcatg
                                                                      960
ttacttattt tgtttccatg tctttttcca aaagaactta ttttttatat tataataaat
                                                                     1020
atcagtggaa aagtaggttt cgttatatag aaattaactt taggctgggt gcagtggctc
aagcctatat ttgggaggcc gaggcaggag gattgcttga actcaggagt tcgaaactag
                                                                     1080
1130
```

10

```
<211> 883
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (19)
<223> n equals a,t,g, or c
<400> 19
gtcaccgtgg gcgtttaant atgatccccg gctcagattc gcagactgca ctgaacttcg
                                                                       . 60
getetacgtt gatgaagaag aagtetgate etgagggtee egegetgete tteeetgaga
                                                                       120
gtgaactttc catccggata ggtagagctg ggcttctttc agacaagagt gagaatggtg
                                                                       180
                                                                       240
aggcatatca gagaaagaag gcggcagcca ctggccttcc agagggtcct gctgtccctg
tgccttctcg agggaatctg gcacagcccg gcggcagcag ctggaggagg atcgcactgc
                                                                       300
tcatcttggc catcactata cacaacgttc cagagggtct cgctgttgga gttggatttg
                                                                       360
gggctataga aaagacggca tctgctacct ttgagagtgc caggaatttg gccattggaa
                                                                       420
tegggateca gaattteece gagggeetgg etgteageet teeettgega ggggeagget
                                                                       480
tetecacety gagagettte tygtatygge agetyagegg catygtyggag cecetygeeg
                                                                       540
gggtctttgg tgcctttgcc gtggtgctgg ctgagcccat cctgccctac gctctggcct
                                                                       600
ttgctgccgg tgccatggtc tacgtggtca tggacgacat catccccgaa gcccagatca
                                                                       660
gtggtaatgg gaaactggca tcctgggcct ccatcctggg atttgtagtg atgatgtcac
                                                                       720
tggacgttgg cctgggctag ggctgagacg cttcggaccc cgggaaaggc catacgaaga
                                                                       780
aacagcagtg gttggcttct atgggacaac aagcttcttt cttcacatta aaactttttt
                                                                       840
ccktcctctc ttcttcaaaa aaaaaaaaaa aaaaaaactc gag
                                                                       883
<210> 20
<211> 989
<212> DNA
<213> Homo sapiens
<400> 20
ctggcttggc tgctatactc ttgcccttca ctgaacctca gttttcctca tctgaatagt
                                                                        60
tgggagactc attectgcct ttetcatgtc cetggctatt tggtaaacca qecaqtaqqa
                                                                       120
agacatcgtg aaatgtatta aagtggtctt agctagacag agtgggcatg ccagggtcag
                                                                       180
cagagattet gaagtetaga ccagtteeet gggtgggeeg ttgteagtee tageagatgg
                                                                       240
ccaggtcagc cctcaggctg gaaattttag ggcagctatt ggtaggtgtc tcctcttgct
                                                                       300
gtgctgagat acggtcaaga tcatacttag gcttttgttg gaagaacata caagacgaga
                                                                       360
gaaaaaaaa gatcatactt aggggctccc ggaatttgct ctgccctagg ttgctgagac
                                                                       420
ctctagaacc tgtgcaggct aaaggaactc agtcggtaga tccgagagag gtggtcaggg
                                                                       480
agaccaggag catgtctaca ctgccagcag acttttgcct cctccccaa gccagcagga
                                                                       540
tggcccaaaa aggctcccc agcagatcat ctttgcagct ccttttttag ctccagtggc
                                                                       600
agcagggatg aggaagggaa agttctatca tttttttcta atttaaaatg acatttaaaa
                                                                       660
tcactagcct agtgggggcc aggtgtggtg gctcatgcct gttgtcccag cactttggga
                                                                       720
ggccaagacg agtggatcgc ttgagctcag gagttaaaga ccagtctggg caacatagcg
                                                                       780
aaacgccgtc tctataaaaa aatacaaaaa ttagctggat gtggtggtgc acacctgtat
                                                                       840
tcttagctgc ttgtggggct aaggcgaaag gatcacttga gcccaggagg tcaaggctgt
                                                                       900
agtgagctgt ttgtgccact gcactctagc ctgggtgaca aagcaaaacc ctgtctcaaa
                                                                       960
aaaaaaaaa aaaaaaaaa ggcggccgc
                                                                       989
```

<210> 21 <211> 495

<212> DNA

<213> Homo sapiens

<400> 21			********	taaggaaggt	atctttcctc	60
ggtggaatgt	agtgaaaacg	agatgetgte	tctwagggac	gaccetgtgc	traggacted	120
ttccgacatt	tgctgactgg	tegatgetag	atgatcttga	gaccccgcgc	ttatctagec	180
agaactacag	gaaaggagca	gtatetetga	attatcgtgt	cactcacac	acaactacta	240
cctgatgact	gtgagcctct	tectectget	tgccacctct	catcacaag	cacctoccag	300
cgactcaggc	tcatgtccca	attcaaggca	gcaagaaggc	cacggagegg	ctacccacat	360
cagatgcccc	tgcaggccat	ctctccaggc	tcaggaacct	aaagaagaat	etacccagac	420
gtggtgctca	cacctgtgat	cccaccactt	tgggaggctg	aggcaagagg	accaccigag	480
cccaggagtt	ggagaccaac	gtaggcaata	cagcaagact	cccatctcta	Caaaaaaaa	495
aaaaaaaac	tcgag					#33
<210> 22						
<211> 2317						
<212> DNA						
<213> Homo	sapiens					
<400> 22	********	aattaatcaa	taaaaattgg	atattgaatt	cataagatgt	60
ccctaaagag	tigaagaact	atactttaa	ttgcttttaa	aaaaattagt	actaactttc	120
taaaatggac	tggatttttg	tttgagttttgg	cctttggtga	accatogaag	tttacccagt	180
aagtgaatta	caaccttaaa	ttogagattet	cecteggega	tototttaa	tracagetet	240
ggtaaggaga	actgtaatgt	ttaggattet	gaataagtac	ttcacaaaca	taaaaacaag	300
taattcaaag	cattgattta	attecaegta	gtctgttata	gestgtstat	acasatagac	360
tggaaaaagt	atgaccccag	tgtcggagat	ggctatgtgt	gcatgcatat	taactcaata	420
acgtatatgt	gcatgttacc	tttacgtaca	tgtggaaaaa	cagttttaat	acctttttaa	480
tgtgtcttgt	atttgcatat	attagagtat	gattttccta	atggtcgagg	gccccccgg	540
tgggccgcag	tttggattta	tgtggtatgt	tgatgaagac	ctagtgaata	gecacageae	600
tcagcttttc	ctctcacaga	ttaatctgcc	agtttctccc	actgtgtatt	gratatacyc	660
gatagaattt	gaggggaaa	ataacacacc	agctaatgat	gaaacgaact	ttttcactt	720
tgtaaggtaa	cgggaccagc	ccaattatat	aactgattga	ggcattgcca	atttagatat	780
aactctttga	atcaagtcat	aattttattt	aataatttt	ctgaggcaga	geeeegeee	840
tgtcgctcag	gctggagtgc	agtggcccaa	tcttggctca	ctgcaacctc	ageeceeag	900
gttcaaggga	ttctcctgcc	tcagcctcct	gagtggctgg	gattacagge	accegecace	960
aagcctagct	atttttttg	tatttttag	tagagacggg	gtttcaccat	gttggeeggg	1020
ctggtctaca	actcctgact	tcaggtgatt	caccggcctc	ggcctcccac	acagetggga	1080
ttacaggtgt	gagccaccgt	geeegaeetg	aatcaagata	ttatttaaaa	gaactgtttg	1140
atgttattta	tttattcatg	tcttcaatag	gtatttacat	gtctgtcttc	ggagatggtg	1200
ttagaggttt	tttttttt	ttttgagacc	gagtcttgct	ctgttgccca	ggctggaggg	1260
cagtggcgcg	atctcggctc	actgcaagct	ccgcctcccg	aattcacgcc	atteteetge	1320
ctcagcctcc	tgagtggctg	ggactacagg	cgcccgccac	cacgcccggc	taatattttg	
catttttagt	ataggtgggg	tttcaccgtg	ttagccagga	tggtgtcgat	ctcctgacct	1380
cgtgattcgc	ccgcttcggc	: ttcctaaagt	gctgggatta	caggcgtgag	ccaccaagcc	1440 1500
caaccaataa	tagagttttg	aacaagacaa	. agctccttgc	aaagctaacc	ttgggaggtt	
tggggaaaac	cacaaagagg	, gtgacaagat	aatttcagat	agggctaagg	gataggaaga	1560
aaatgaagat	gggacatgtt	: aagtaataat	. aatagatgac	atttgaacac	tgtgccagtc	1620
actctqtcqa	ttgttttaca	tataccttgt	. gacaacccta	. agaggtaggc	accgttatta	1680
cagacatttt	acaggtgaaa	a cagggacaca	ggaaaagtaa	. gtaacatgcc	cagctgttga	1740
atcaaggcag	cccgggacca	gagtccactc	: tctgagaaat	ggcatttggg	caggatcgag	1800
aatgaggagg	agatccagto	tggcagaggc	cctggggagg	aatgggctgg	aagtattcga	1860
дааасссааа	agtaagagto	gcctgagtgg	cccagagagg	r ttgatgggag	gcaggagtca	1920
gagcgtgtct	accetegtag	gccattggag	gcttacatgt	. ggaagggaca	ggttctgatt	1980
cctatttaaa	aagagggctg	g ctgtccgggc	gcggtggctc	: acgcctgtag	tcccagcact	2040
ttgggaggg	gagggggg	gatcgcgagg	tcgggagatc	gagaccatco	: tggctaacac	2100
	J-33-335-					

PCT/US99/05804

WO 99/47540

12

2160 cccgtctcta ctaaaaatac aaaaacaaaa ttagcccggc atggtggcgg gcgcctgtgg tctcagctgc tcgggaggct gaggcgaaag aatggcgtga acctggaagg cggagcttgc 2220 2280 agtgagccgg gattgtgcca ctgcactcca gcttgggcga cagagcgaga ctccatctcc 2317 aaaaaggaat tcgatatcaa gcttatcgat accgtcg <210> 23 <211> 1726 <212> DNA <213> Homo sapiens <400> 23 ctttttggct ctcattttga atttttcaag agctcatgtt ctttgtcttc attaaaaaaa 60 aaaagttott atatogtgta tgaatgtoat togggatato tatacataca tgcacataco 120 tcatttttat tgcatttcac tttattgcac tttgcaaagt gacatttttt acagattcaa 180 ggtttggcaa ccctatgtcc ataagtctgt cagcaccatt tttctaacat atgtgctcat 240 tttgcctctc tgtcacattt ttttttttc ctgagacagg gtcttgctct gtcacccagg 300 ctggaatgca gtggtgtaat tatgcctcac tgcggccttg acctcttggg ctcaagggat 360 420 cctcttgcct gcgcttcttg agtagctgag actacagatg tacaccacca cacaccage taagttttaa attttttat agagatgggg tttccctatg ttgcccaagc tgctctcgaa 480 ctcctgggct taagtgatcc tcccacctca gcctttcaaa gtgctgggat tacaggcatg 540 agccacagca cctggtctct gtgtcacgtt ataataattc tgccaatatt ccagatttcg 600 tcattattaa atctgttatg gtgatctgtg atcagcgaac tctgatgtta ctgtctaatt 660 720 gctttgggat gcagtgaacc cgtcagagtc atccatgagg gttggaatca acttcttcca aaatcctgtt aatcaagagt gaacttaatc gattaatgtt gtgtatgttc tgactgctcc 780 accaatctgt gggtccccta tcactctccc tctcctcagg cctccctatt ccctgagaca 840 caataatatt gaaattagac caattaataa ccctgcaatg tgaaaggaag aagttacatg 900 tctctcactt taaatcaaaa gctagaaatt attaagctta gtgaggaggg catttcgaaa 960 gctgagagag gctgaaagct aggccggttg tgccaaatag ctagccaagt tgtgaatgca 1020 taggaaaagt tcttgaagga aattaaaatt gttactccag tgaacacaca aatgttaagt 1080 aagcaaaaca gccttattgc ttatggaaag aaagtctgaa tggtctgaat agaagatcac 1140 1200 atcagccaaa acatgtcctt aagccaaagc ctaatctata atagatcagg ccctaagtct cttccattcc ttgaaggcac agagggtga agaagttgca gaagaaaagt tggaagctag 1260 ccaaccttgg tttgtgcagt ttaaggaaaa aagccatctc cataacatgg aagtgcaaga 1320 tgaagcagca ggcactagtg gggaagctgc agcaagttat acagaaaatc tagctaatga 1380 1440 tgagggtggc tacactaaaa acagattttc aatggagaca aaacaccctt ctattggaag 1500 aagatgccct ctaaagcttt cataggtaga gagaggtcag tgcatgggct tgaaagaaca 1560 ggctgattct cttgctagag gccaatgaag ccagtgagtt taagttgaag ccagtgctaa 1620 tttatcattc tgaaaattgt agggccctta agcattatgc taaatctact ctgcctgtgc 1680 tctagaaatg gaatagcaaa gcacgaatga cagcacatct gtttacaaca tgatgtgctg aatattttaa gootatattt gagacotact gotcaaaaaa aaaaaa 1726 <210> 24 <211> 529 <212> DNA <213> Homo sapiens <400> 24 acgcgtccga ttacttacgt gctcctggct gggatggcac tgggcattca gaaaaggttc 60 120 tecceggagg tgctgggect gtgtgcaage acagegetgg tgtgggtggt gatggaggtg .180 ctggccctgc tcctgggcct ctacctggcc accgtgcgca gtgacctgag cacctttcac 240 ctgctggcct acagtggcta caaatacgtg ggaatgatcc tcagtgtgct cacggggctg 300 ctgttcggca gcgatggcta ctacgtggcg ctggcctgga cctcatcggc gctcatgtac 360 ttcattgtgc gctctttgcg gacagcagcc ctgggccccg acagcatggg gggccccgtc

```
420
ccccggcage gtctccagct ctacctgact ctggggagctg cagccttcca gcccctcate
                                                                      480
atatactggc tgactttcca cctggtccgg tgaccccctg gccccagatg gcactgagtt
529
<210> 25
<211> 1755
<212> DNA
<213> Homo sapiens
<400> 25
                                                                      60
qqcacqagcc tcacaqcqcc tctqctqqag ttcctqctqq ccttqtactt cctctttqct
                                                                      120
gatgccatgc agctgaatga caagtggcag ggcttgtgct ggcccatgat ggacttcctg
                                                                      180
cactatata cogoggeest catetactit getatetesa teaeggeeat egecaagtae
                                                                      240
teggatgggg cttccaaagc cgctgggggg tetgtgcctg acactegggc tgtttgtcca
                                                                      300
agcagatctg aaatgggccg tgagctgggg gcagcagcct cccgggagca gggagtcagc
                                                                      360
cctgtgatgc atcccatcca ccctgtccac aggtgtttgg cttctttgct accatcgtgt
                                                                      420
ttgcaactgr tttctacctg atctttaacg acgtggccaa attcctcaaa caaggggact
                                                                      480
ctgcagatga gaccacagcc cacaagacag aagaagagaa ttccgactcg gactctgact
                                                                      540
gaaggeetge gggtgeettg geaacetgag ceacacagge etceacecet gegeetcaca
                                                                      600
ggggtcgctg gcgttggagc ggaggcctgg acttctgagt tgcagagggg gctgcggaca
                                                                      660
cagcaggece cetacagect caggttetge etgageceag cetaceagge ttgccctca
gctcagcact gttgaccacg ctgcgtatga gggcatcttg ggtatcccac tccttctccc
                                                                      720
                                                                     780
catttctgtc ccacaggcct tcagcccttt aacgtctctg ccaaaaacca gcacaaggag
                                                                      840
acaaagcaga gccttgtctg tatctgggca gcaggtgttc catgctgcta ggtggcgggg
                                                                     900
qtcqqqqqtc ttctqtttca ctaacaggaa caaagacaga aaccatgaca gggctgcccc
                                                                     960
qccaqqccc qqtgggtttg tctgcacttg gtgctcctgc ccacaccagc cactttggtg
                                                                     1020
acaatgaccc ttccaagaat ctttggttca aggagcacca gttccctctt cattcttgaa
                                                                     1080
qcaqqaqaa attqaccttt gccttgtcgc ccaggaagtg gggctcggca cccataacta
                                                                     1140
acacetecca ecettggaaa ecatgtette tgggggtgag atgaceatte tgggtetaag
                                                                     1200
actgtttcaa agaagagete atagactgac tggtccagaa gacagagggt acaacagtgg
                                                                     1260
catcacagtg acagtgtcat ggggagctgg gcgggcccag ccaaaccctc cttcttccta
gageccagec ageaggeagg agtteetgga ceetcaggac agtgaactte cagacetcag
                                                                    1320
                                                                    1380
ggcaggtcta tgggccactg caggagatga gaccagcctt ctgtgttcac ctaacgattt
atactgtgta tctgtctttg atggaatttt gtaacttttt atatttttt atgcaaaagc
                                                                    1440
agettettaa cagatggeat tttetgtgae tetaggeete acaaaagage cagagttetg
                                                                    1500
gacccatgtt tggagcattt gtagccttat tctcttgcgt gtgaatctct taccctgaaa
                                                                    1560
aaaagccata atgaattaag ccagactgac cacttgcttg gagtgtgtgc ttgaaaaaaac
                                                                     1620
cagagcaata ctgttgggta ttgtatcagg cttcagtaca aactggtaac accaatgtgg
                                                                     1680
atcctgacag ctttcagttt tagcaaaaat acacgtgaaa tctgactacc atttaaaaaa
                                                                     1740
                                                                     1755
aaaaaaaaa aaaaa
<210> 26
<211> 1751
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (1520)
<223> n equals a,t,g, or c
<220>
<221> SITE
```

14

<222> (1557)

```
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (1689)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (1729)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (1735)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (1741)
<223> n equals a,t,g, or c
<400> 26
gggtgcagcc tgatggcgca ggaggtagac acggcacagg gcgccgagat gcggcggggc
                                                                        60
gegggegegg cteggggaeg egetteetgg tgetgggeee tggegetget ttggetegeg
                                                                       120
gtggttccgg gctggtcccg ggtctcgggc atcccctccc ggcqccactq qccqqtqccc
tacaagcgct ttgacttccg tccaaaacct gatccttatt gtcaagctaa gtatactttc
tgtccaactg gctcacctat cccagttatg gagggtgatg atgacattga agtttttcga
ttacaagccc cagtatggga atttaaatat ggagacctcc tgggacactt gaaaattatg
                                                                       360
catgatgcca ttggattcag aagtacatta actggcaaga actacacaat ggaatggtat
                                                                       420
gaacttttcc aacttggcaa ctgtacattt ccccatctcc gacctgaaat ggatgccct
                                                                       480
ttctggtgta atcaaggcgc tgcctgcttt tttgagggaa ttgatgatgt tcactggaag
                                                                       540
gaaaatggga cattagttca agtagcaact atatcaggaa acatgttcaa ccaaatggca
                                                                       600
aagtgggtga aacaggacaa tgaaacagga atttattatg agacatggaa tgtaaaagcc
                                                                       660
agcccagaaa agggggcaga gacatggttt gattcctacg actgttccaa atttgtgtta
                                                                       720
aggaccttta acaagttggc tgaatttgga gcagagttca agaacataga amccaactat
                                                                       780
acargaatat ttctttacag tggagaacct acttatctgg gaaatgaaac atctgtttt
                                                                       840
gggccaacag gaaacaagac tcttggttta gccataaaaa gattttatta ccccttcaaa
                                                                       900
ccacatttgc caactaaaga atttctgttg agtctcttgc aaatttttga tgcagtgatt
                                                                       960
gtgcacaaac agttctattt gttttataat tttgaatatt ggtttttacc tatgaaattc
                                                                      1020
ccttttatta aaataacata tgaagaaatc cctttaccta tcagaaacaa aacactctct
                                                                      1080
ggtttataaa acaccttaat tctactgctc ttttttctcc aatcaccagc atctgttttt
                                                                      1140
cagggggtga ttttactttt gtgaattcct tagcctttct tccttggtgc ataaagttaa
                                                                      1200
aatgcacatc agcagaattg ctgcatatta acatctcagg actcttctct tgtaaagaag
                                                                      1260
ctgaaattcg tactatattg gccaaagtga gcgagttagg tgatcttggt ttcaatttcc
                                                                      1320
gagcettigt taatatggag aattatggtt catateagtt atgtaggace tttggaceca
                                                                      1380
gggtcctaca gatagatatg gtgtgcccag attttaaaaa taccttcaaa aataaaaaat
                                                                      1440
acattcagtg acattttcat ggtgggagct cttctttctg atatggcagt tacacttttt
                                                                      1500
cacttaagtg ctttagtttn agactaactt tacaacttct ataacttttg ggaaccnagt
                                                                      1560
ttagtatagt ctgattacat tccattcacc taactttagg cattcggttt agacaccata
                                                                      1620
actggrgkgr atkgkgcytc cyagratgtg ggcaaatccc agtggttaac accatatttc
                                                                      1680
tgggctggng attttgggga ctagctaggt aaacgggctt ggtggttcnt ttaancatac
                                                                      1740
ntaaccacca c
                                                                      1751
```

```
<210> 27
<211> 1212
<212> DNA
<213> Homo sapiens
<400> 27
qccaaqcttg gcacgargtt ggtggcggcg tccggaggtg ctggtttgtt ctcggtgaac
                                                                       60
qgcgcgcggg gtctctcctg agtgcgagct acgggacctt cgccatgccg gggatggtac
                                                                      120
tottoggcog gogotgggco atogcoagog acgacttggt ottoccaggg ttottogago
                                                                      180
                                                                      240
tggtcgtgcg agtgctgtgg tggattggca ttctgacgtt gtatctcatg cacagaggaa
                                                                      300
agctggactg tgctggtgga gccttgctca gcagttactt gatcgtcctc atgattctcc
                                                                      360
tggcagttgt catatgtact gtgtcagcca tcatgtgtgt cagcatgaga ggaacgattt
gtaaccctgg accgcggaag tctatgtcta agctgcttta catccgcctg gcgctgtttt
                                                                      420
                                                                      480
ttccagagat ggtctgggcc tctctggggg ctgcctgggt ggcagatggt gttcagtgcg
acaggacagt tgtaaacggc atcatcgcaa ccgtcgtggt cagttggatc atcatcgctg
                                                                      540
                                                                      600
ccacagtggt ttccattatc attgtctttg accetcttgg ggggaaaatg gctccatatt
                                                                      660
cctctgccgg ccccagccac ctggatagtc atgattcaag ccagttactt aatggcctca
                                                                      720
agacagcagc tacaagcgtg tgggaaacca gaatcaagct cttgtgctgt tgcattggga
                                                                      780
aagacgacca tactcgggtt gcttyttcga gtacggcaga gcttttctca acctactttt
                                                                      840
cagacacaga totggtgccc agogacattg cggcgggcct cgccctgctt catcagcaac
aggacaatat caggaacaac caagacctgc ccaggtggtc tgccatgccc cagggagctc
                                                                     900
                                                                     960
ccaggaagct gatctggatg cagaattaga aaactgccat cattacatgc agtttgcagc
                                                                     1020
ageggeetat gggtggseee tetacateta cagaaaceee etcaegggge tgtgeaggay
                                                                     1080
tggtggtgac tgaaattagc tggacatggt tgcacacacc tgtaatcaca gctactcggg
                                                                     1140
aggttgaggc gggagaatcg cttgaaccag ggagttggag gttgcagtga gtggagatca
                                                                     1200
1212
aaaaaactcg ag
<210> 28
<211> 1112
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (1105)
<223> n equals a,t,g, or c
<400> 28
qqcacqaqca aacatccagg agtgtgcacc ggtcatgcaa ggtgttttgt ttggctttgt
                                                                      60
                                                                      120
ctqqcttttt agttttttgt ggcaggagaa taaatctagt gcctctccct ccacattagc
caaaagtgga agtccctgtc cagtcagcat tccttggatg cctggtgtat tagtccgttt
                                                                      180
                                                                      240
tttcacactg ctataataaa aagaactgcc caagactggg taactaataa aggaaagagg
                                                                      300
tttaattgac tcacacttct gcatgtttgg gaggcctcag gaaagttaca atcaggcaga
aggtgaagtg cgttcgtctt aatggcggca ggtgagacag tgtgtaggat aaactgtcaa
                                                                     360
                                                                      420
acacttataa aaccatcata gctcatgaga cctcattcac tgtcacgaga acagcatggg
                                                                      480
ggaaccgccc ccatgatcta atcacctccc actaggtccc tccctccacc tgtggggatt
                                                                     540
atgaggatta caattcaaga tgagatttgg gcaggggcac cgagccaaac catatcacct
                                                                     600
tatatgtgcc cagtgttgac ctaggcgctg ggatgcagaa acaaacacga catgggctgt
                                                                     660
gccttgggga gctcacactc ttgctggaga agcatgctga ttcctaaata agaaatgcta
                                                                     720
tgtgctgtgt acagagtacc atggaaggca ggatgaactc tttgggagga agaagcaagg
aaaqctttaq agagttgctg gcttttgagg gatggagcag gcattttcta catggggaaa
                                                                     780
qtqtaqqaaa qagtattcca ggcagagtgg agagcaagag caaaggcgga gaagcctgtg
                                                                     840
```

PCT/US99/05804

caagtgactg atggagtgag taccaagctg	cttgccgggc gaagtgtggt caacccatgg aggagcacaa gctctgatag	ttttgttctg agaaatagag gatttggaac	ggactagtag gctcggggtc ataataccaa	gtagaacaga agctgataca	aagagttggg aggcgttgta	900 960 1020 1080 1112
<210> 29 <211> 748 <212> DNA <213> Homo	sapiens					
gaatgagagt ttttagecat tggtgtkgaa tctgttcaga aaagaaagct tctatcagat ttgataaatg agcactttgg aacatggaga cctgtaatcc aggttgcagt	aaactgttt tgctgttgct cctaataagt catctttccc tcttttgcc ttttttatta aggtcttttg gctttaaaat gaggccaagg aaccccatct cagctacttg gagccgaaat aaaaaaaaaa	ccacggcctc gttagtggct catgtttatt gtttttgttt ttgagttgta caaatatttt aataatctgg gcagatcatc ctactaaaaa agaggctgag cgtgccactg	accagcattt atcattgttt tgtcatctgc gcttgcatgt atagtgcttg ccccaatctg ccgggcgcag tgaggtcggg tataaaatta acaggagaat	ggtggtgtca tcatttgcaa atatcttctt ttgtttgtgt tatagtgtgg tggactgtct tggctcatgc agttcgagac gtcgggcgtg ctcttgaacc	gtgtcttgga ttctcttaca cggccagtta ttgattttt ataacagttc tctcattctt ctgtaattcc cagcctgacc gaggcacatg cgggaggtgg	60 120 180 240 300 360 420 480 540 600 660 720 748
<210> 30 <211> 778 <212> DNA			,			
<213> HOMO	sapiens					
<400> 30 ggaactaaaa cagggagaat ctctaacctt cttgctctcc ggcaccacct ttccagtgca ttctgagcac tcatttccta tccttttcct tgaggcatta tcagatcatt acttctagct	agctttgtgt gaccataggg agggtaggga tcgtgtctgc ctccaggagg tccccaatcc aagtctgtgc gacaaccaag cataccaact cctcattgaa tctcaagctg tgccacagac gacatatcag	cactaagtaa gtgctaagga tgttttttc attggaaagg caccatctta gtaatgcaac attctcaata tgactatgtt ctttagctgc gctcccagct acaattctac	ggctgggatt tctggggaaa ccttggtgta aggagtaaag cctggaatat tgactctctt agtcccagtc tcactgaaac ctgcatgagt tagggcaaag tccaaagtca	ggatcagcag ccatgggctg ctatacagag gattctgatt aaggctgcct actttttct tcatcacaaa ccacaggtct cctctgtcct agaatctcca gcctggcata	aaatccaacc ggaagctgct gccagatgit tgattgatga tgtacccctt tagtaactga tattaatatt tgggacagaa caagtctttc tgatgtgctg gtaacattga	60 120 180 240 300 360 420 480 540 600 660 720 778

<400> 31

120

acgagctaat gattcttgct gaagatggcc agagtaatca gagtaattaa tttggggaat

```
ggtgaaggat aaggaacctg atcagatatc aagggtgggg gtactcttgc taaactgact
cagtagggtc cttgctaaaa ctggttttta caagagagag cacaggtagg tctacaagaa
                                                                     180
                                                                     240
gattcgggag actgactaaa gtttcatcaa gaatctttgt caggagacat gacctttatg
                                                                     300
atacttaagt ttttcttttt atgeggtttt gttttaaaca ggctaatagc tegtcagett
gctaaaatcc atgctattca tgcacacaat ggctggatcc ccaaatctaa tctttggcta
                                                                     360
aagatgggaa agtatttctc tctcattccc acaggatttg cagatgaaga cattaataaa
                                                                      420
aggtaaaatt atttttacat ttgaaattat gttttacatt gctttgctct atgatagggt
                                                                      480
ttcaaaggta attgtaaagt ttcattgtat aaaatctggt tttctttctt tgcattgaag
                                                                     540
taaagtaagc attgattctt tggctgtcag atagcactac agaaataact gcctctccat
                                                                     600
                                                                      660
cccctcagt gtcccctccc aaaaaatatg ccctacaaca gcaaggggca gaagtggaag
                                                                      720
taggggaaca ccacttaaaa ataaagaggt aggtggtaat ggtgaagcaa ggatttttct
                                                                     780
tgacttttta agatactgca gggttgagag gcaagtctag tattatatag aataagagca
                                                                     840
caagagcctg gagccaaact gagattaaat cttagtcctg ctagttaaca tttctgttgt
                                                                     900
gtaactcatt caactaggga acttaaccta actgtttcct tatctataaa atggaaatta
                                                                     960
cagtagtagt atattaacca tatagagttt ttgtgaggat tgagatagta tatgtaaagt
                                                                     1020
totttaaaac agtgootggt catcacttaa gtgttgaggt agctgctgtt ttaaaaatta
                                                                    1080
ctattgttat tcaaagaagg ttatttgagt tatatttttt ccctaggctc tccaaggtat
                                                                    1140
actttaaatc cttgaggtta atgattcttt ggaaaagctg gaggtgtgct gtggtaaata
                                                                    1200
ataggaagaa aacccttgcc ccaatagaaa ataatacaac tagaacataa aacacaatta
                                                                    1260
aaatattaaa tacattttgt atttctttta ttccattttg tttgttcatg atttaagttt
                                                                    1320
1324
aaaa
<210> 32
<211> 739
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (732)
<223> n equals a,t,g, or c
<400> 32
ggcacgagga caggatcctg gtttgggtac cttagtttaa tagaaaccca ggtggaaacc
                                                                       60
tagattgcag ggcatttgtt tgagactatt tagccacagc agggcaagca ggaagatgca
                                                                      120
gcctgtcact gctaactcct tagtattaaa actgtcaaac atgggaggta actgctgatg
                                                                      180
catttttcag ttgatagttt atatactttc tctgaaggat cctaatgata gttaaccatt
                                                                      240
tctcattttt attttgctgg attgttttct gttttttgct tcagcattct tgcttttgct
                                                                      300
                                                                      360
gtgcttactt ttggagtttt gattccctgt gtcactgttt tctttcgcat acacctctca
ggtttacaca gtaaacaatg tgaatgtgat caccaaaata cgcacagaac atctgaccga
                                                                      420
                                                                      480
ggaggaaaaa aagagatata aaggtaatca ccaccacct cccacctcct gttttgttgt
                                                                      540
tattttttaa gcctagaggg aactctttgt tggctctgtt aagtttaggg ttaatgtgat
                                                                      600
tgggttgtgt taagcctaac cctaacttct wetctetete tetettttt ttttttttg
                                                                      660
aggcaaagtc tegecetgte acceaggetg gagtgeagtg geacgaeett ggeteactge
                                                                      720
aacctctgcc tcctgaggca ggagaatcac ttgaacctgg caggcggagg ttttggtgag
                                                                      739
ctgaggtcgt gncattgca
```

<210> 33

<211> 1462

<212> DNA

<213> Homo sapiens

<400> 33						
ggccatcggc g	gggcagtcg	cgggatgcgc	ccgggagcca	cagcctgagc	tttagcccat	60
gaggaggatg t	gaccgggac	tgagtcagga	gccctctgga	agcatggaga	ctgtggtgat	120
tgttgccata g	gtgtgctgg	ccaccatctt	tctggcttcg	tttgcagcct	tggtgctggt	180
ttgcaggcag c	gctactgcc	ggccgcgaga	cctgctgcag	cgctatgatt	ctaagcccat	240
tgtggacctc a	ttggtgcca	tggagaccca	gtctgagccc	tctgagttag	aactggacga	300
tgtcgttatc a	ccaaccccc	acattgaggc	cattctggag	aatgaagact	ggatcgaaga	360
tgcctcgggt c	tcatgtccc	actgcattgc	catcttgaag	atttgtcaca	ctctgacaga	420
gaagcttgtt g	ccatgacaa	tgggctctgg	ggccaagatg	aagacttcag	ccagtgtcag	480
cgacatcatt g	tggtggcca	agcggatcag	ccccagggtg	gatgatgttg	tgaagtcgat	540
gtaccctccg t	tggacccca	aactcctgga	cgcacggacg	actgccctgc	tcctgtctgt	600
cagtcacctg g	tgctggtga	caaggaatgc	ctgccatctg	acgggaggcc	tggactggat	660
tgaccagtct c						720
ttctgagcca g	ataaaggcc	tcccaggccc	tgaaggcttc	ctgcaggagc	agtctgcaat	780
ttagtgccta ca						840
agcttagcct to						900
tgtgtgtgca ta						960
ctgtggctgg t						1020
aagaggagta t				_		1080
gaaaattgag c	caccgtcta	agaaatcaag	aggtttcaca	ttaaaattag	aatttctggc	1140
ctctctcgat c	ggtcagaat	gtgtggcaat	tctgatctgc	attttcagaa	gaggacaatc	1200
aattgaaact a				-		1260
tttcatttat t						1320
caggtttggg t						1380
ctgcaatttt a	cctagctac	cactaggtgg	atagtaaatt	tatacttatg	tttccctcaa	1440
aaaaaaaaaa a	aaaaactcg	ag				1462
.010. 24					•	
<210> 34						
<211> 2815						
<211> 2815 <212> DNA						
<211> 2815	apiens					
<211> 2815 <212> DNA <213> Homo sa	apiens					
<211> 2815 <212> DNA <213> Homo seconds		ctgatagaga	ctatagttcg	agagttettg	cccaccagtt	60
<211> 2815 <212> DNA <213> Homo so <400> 34 gggtcctgga gg	tgeeetegg					60 120
<211> 2815 <212> DNA <213> Homo se <400> 34 gggtcctgga ggggtctcctgt gg	tgccctcgg ggggcaggg	cctaccccta	gtctatacaa	agtaccctgt	gctactgcca	120
<211> 2815 <212> DNA <213> Homo sa <400> 34 gggtcctgga gg ggtctcctgt gg tgaaactact to	tgecetegg ggggeaggg egtgteetg	cctaccccta gcctcagctg	gtctatacaa ggaggaatat	agtaccctgt tgctgcccgg	gctactgcca ctgttgagca	120 180
<211> 2815 <212> DNA <213> Homo sa <400> 34 gggtcctgga gg ggtctcctgt gg tgaaactact to gctttgatct co	tgeeetegg ggggeaggg egtgteetg eggageege	cctaccccta gcctcagctg ctgtgccgca	gtctatacaa ggaggaatat tcatagctga	agtaccctgt tgctgcccgg ggctccccaa	gctactgcca ctgttgagca gaactggcct	120 180 240
<211> 2815 <212> DNA <213> Homo sa <400> 34 gggtcctgga gg ggtctcctgt gg tgaaactact to gctttgatct cc tgccccaga gg	tgccctcgg ggggcaggg cgtgtcctg cggagccgc gaagctgag	cctaccccta gcctcagctg ctgtgccgca atgctgagca	gtctatacaa ggaggaatat tcatagctga ccgaggccct	agtaccetgt tgctgcccgg ggctccccaa ccgtctgtgg	gctactgcca ctgttgagca gaactggcct gctgtggctg	120 180 240 300
<211> 2815 <212> DNA <213> Homo sa <400> 34 gggtcctgga gg ggtctcctgt gg tgaaactact to gctttgatct tgccccaga gg cctcctatgg cc	tgccctcgg ggggcaggg cgtgtcctg cggagccgc gaagctgag cagggcggt	cctaccccta gcctcagctg ctgtgccgca atgctgagca tacctttaca	gtctatacaa ggaggaatat tcatagctga ccgaggccct gggagctcta	agtaccctgt tgctgcccgg ggctccccaa ccgtctgtgg cccagtgctg	gctactgcca ctgttgagca gaactggcct gctgtggctg atgcgggcct	120 180 240 300 360
<211> 2815 <212> DNA <213> Homo sa <400> 34 gggtcctgga gg ggtctcctgt gg tgaaactact ta gctttgatct ca tgccccaga gg cctcctatgg tgcaggtggt ga	tgccctcgg ggggcaggg cgtgtcctg cggagccgc gaagctgag cagggcggt ccgcgggag	cctaccccta gcctcagctg ctgtgccgca atgctgagca tacctttaca ctcagcaccc	gtctatacaa ggaggaatat tcatagctga ccgaggccct gggagctcta acccacctca	agtaccctgt tgctgcccgg ggctccccaa ccgtctgtgg cccagtgctg acccctgtcc	gctactgcca ctgttgagca gaactggcct gctgtggctg atgcgggcct atgcagcgga	120 180 240 300 360 420
<211> 2815 <212> DNA <213> Homo sa <400> 34 gggtcctgga gg ggtctcctgt gg tgaaactact to gctttgatct co tgccccaga gg cctcctatgg cc tgcaggtggt gg tagcctcact gg	tgccctcgg ggggcaggg cgtgtcctg cggagccgc gaagctgag cagggcggt ccgcgggag ctcactct	cctaccccta gcctcagctg ctgtgccgca atgctgagca tacctttaca ctcagcaccc ctcacccagc	gtctatacaa ggaggaatat tcatagctga ccgaggccct gggagctcta acccacctca taaccctggc	agtaccetgt tgctgcecgg ggctccccaa ccgtctgtgg cccagtgctg acccetgtcc agccggcagt	gctactgcca ctgttgagca gaactggcct gctgtggctg atgcgggcct atgcagcgga acccctgctg	120 180 240 300 360 420 480
<211> 2815 <212> DNA <213> Homo sa <400> 34 gggtcctgga gg ggtctcctgt gg tgaaactact tg gctttgatct cg tgccccaga gg cctcctatgg cg tgcaggtggt gg tagcctcact gg aaaccatcag tg aaaccatcag tg	tgccctcgg ggggcaggg cgtgtcctg cggagccgc gaagctgag cagggcggt ccgcgggag ctcactctc	cctaccccta gcctcagctg ctgtgccgca atgctgagca tacctttaca ctcagcaccc ctcacccagc gaggccagcc	gtctatacaa ggaggaatat tcatagctga ccgaggccct gggagctcta acccacctca taaccctggc tctcggccac	agtaccetgt tgctgcecgg ggctccccaa ccgtctgtgg cccagtgctg acccctgtcc agccggcagt cccttcctta	gctactgcca ctgttgagca gaactggcct gctgtggctg atgcgggcct atgcagcgga acccctgctg gtcacttgga	120 180 240 300 360 420 480 540
<211> 2815 <212> DNA <213> Homo sa <400> 34 gggtcctgga gg ggtctcctgt gg tgaaactact tg gctttgatct cg tgccccaga gg cctcctatgg cg tgcaggtggt gg tagcctcact gg aaccatcag tg cacaggtgtc tg cacaggtgtc tg cacaggtgtc tg cacaggtgtc tg	tgccctcgg ggggcaggg cgtgtcctg cggagccgc gaagctgag cagggcggt ccgcgggag ctcactctc gattctgct	cctaccccta gcctcagctg ctgtgccgca atgctgagca tacctttaca ctcagcaccc ctcacccagc gaggccagcc cctcttgttg	gtctatacaa ggaggaatat tcatagctga ccgaggccct gggagctcta acccacctca taaccctggc tctcggccac agccgtgtct	agtaccetgt tgctgcecgg ggctccccaa ccgtctgtgg cccagtgctg acccctgtcc agccggcagt cccttcctta aaggcagacc	gctactgcca ctgttgagca gaactggcct gctgtggctg atgcgggcct atgcagcgga acccctgctg gtcacttgga ttgaagttgc	120 180 240 300 360 420 480 540
<211> 2815 <212> DNA <213> Homo sa <400> 34 gggtcctgga gg ggtctcctgt gg tgaaactact tg cttgatct cc tgccccaga gg cctcctatgg cc tgcaggtggt gg tagcctcact gg aaccatcag tg cacaggtgtc tg tgtccagacc tg tgccagacc tg tgccagacc tg	tgccctcgg ggggcaggg cgtgtcctg cggagccgc gaagctgag cagggcggt ccgcgggag ctcactctc gattctgct gggctccag	cctaccccta gcctcagctg ctgtgccgca atgctgagca tacctttaca ctcagcaccc ctcacccagc gaggccagcc cctcttgttg agagccgtgg	gtctatacaa ggaggaatat tcatagctga ccgaggccct gggagctcta acccacctca taaccctggc tctcggccac agccgtgtct gcccagtgcc	agtaccetgt tgctgcecgg ggctccccaa ccgtctgtgg cccagtgctg acccctgtcc agccggcagt cccttcctta aaggcagacc cgttgcctgc	gctactgcca ctgttgagca gaactggcct gctgtggctg atgcgggcct atgcagcgga accctgctg gtcacttgga ttgaagttgc ctgttgttcc	120 180 240 300 360 420 480 540 600
<pre><211> 2815 <212> DNA <213> Homo sa <400> 34 gggtcctgga gg ggtctcctgt gg tgaaactact tg gctttgatct cg tgccccaga gg cctcctatgg cg tgcaggtggt gg tagcctcact gg aaaccatcag tg cacaggtgtc tg tgtccagacc tg tgtccagacc tg tgggagccta cg</pre>	tgccctcgg ggggcaggg cgtgtcctg cggagccgc gaagctgag cagggggt ccgcgggag ctcactctc gattctgct gggctccag gagatgtgg taccaggcc	cctaccccta gcctcagctg ctgtgccgca atgctgagca tacctttaca ctcagcaccc ctcacccagc gaggccagcc cctcttgttg agagccgtgg tggagccagc	gtctatacaa ggaggaatat tcatagctga ccgaggccct gggagctcta acccacctca taaccctggc tctcggccac agccgtgtct gcccagtgcc aaccaagctc	agtaccctgt tgctgcccgg ggctccccaa ccgtctgtgg cccagtgctg acccctgtcc agccggcagt cccttcctta aaggcagacc cgttgcctgc atgcccggag	gctactgcca ctgttgagca gaactggcct gctgtggctg atgcgggcct atgcagcgga acccctgctg gtcacttgga ttgaagttgc ctgttgtcc	120 180 240 300 360 420 480 540
<pre><211> 2815 <212> DNA <213> Homo sa <400> 34 gggtcctgga gg ggtctcctgt gg tgaaactact tg gctttgatct cg tgccccaga gg cctcctatgg cg tgcaggtggt gg tagcctcact gg aaaccatcag tg cacaggtgt tg tgtccagacc tg tgtccagacc tg tgggagccta cg aggacatgga gg</pre>	tgccctcgg ggggcaggg cgtgtcctg cggagccgc gaagctgag caggggggt ccgcgggag ctcactctc gattctgct gggctccag gagatgtgg taccaggcc cgcctgtca	cctaccccta gcctcagctg ctgtgccgca atgctgagca tacctttaca ctcagcaccc ctcacccagc gaggccagcc cctcttgttg agagccgtgg tggagccagc	gtctatacaa ggaggaatat tcatagctga ccgaggccct gggagctcta acccacctca taaccctggc tctcggccac agccgtgtct gcccagtgcc aaccaagctc gctgccactg	agtaccctgt tgctgcccgg ggctccccaa ccgtctgtgg cccagtgctg acccctgtcc agccggcagt cccttcctta aaggcagacc cgttgcctgc atgcccggag ctgagtcagc	gctactgcca ctgttgagca gaactggcct gctgtggctg atgcgggcct atgcagcgga acccctgctg gtcacttgga ttgaagttgc ctgttgttcc gattggctcc ccacactggg	120 180 240 300 360 420 480 540 600 660 720 780
<pre><211> 2815 <212> DNA <213> Homo sa <400> 34 gggtcctgga gg ggtctcctgt gg tgaaactact tg gcttgatc cc tgccccaga gg cctcctatgg cc tgcaggtggt gg tagcctcact gg aaaccatcag tg cacaggtgt tg tgtcagacc tg tgtcagacc tg tgggagccta cg aggacatgga gg cagcctgtgg gg cagcctgtgg gg</pre>	tgccctcgg ggggcaggg cgtgtcctg cggagctgag caggggggt ccgcgggag ctcactctc gattctgct gggctccag gagatgtgg taccaggcc ccgctgtca attccctta	cctaccccta gcctcagctg ctgtgccgca atgctgagca tacctttaca ctcagcaccc ctcacccagc gaggccagcc cctcttgttg agagccgtgg tggagccagc gagagctgct ggcactgct	gtctatacaa ggaggaatat tcatagctga ccgaggccct gggagctcta acccacctca taaccctggc tctcggccac agccgtgtct gcccagtgcc aaccaagctc gctgccactg ccttctctgc	agtaccctgt tgctgcccgg ggctccccaa ccgtctgtgg cccagtgctg acccctgtcc agccggcagt cccttcctta aaggcagacc cgttgcctgc atgcccggag ctgagtcagc aacccgctgt	gctactgcca ctgttgagca gaactggcct gctgtggctg atgcgggcct atgcagcgga acccctgctg gtcacttgga ttgaagttgc ctgttgttcc gattggctcc ccacactggg cctgtgtcc	120 180 240 300 360 420 480 540 600 660 720 780 840
<pre><211> 2815 <212> DNA <213> Homo sa <400> 34 gggtcctgga gg ggtctcctgt gg tgaaactact tg gcttgatct cc tgccccaga gg cctcctatgg cc tgcaggtggt gg tagcctcact gg aaaccatcag tg cacaggtgt tg tgtccagacc tg tgtcagacc tg tgggagccta cg aggacatgga gg agccttgaa gg agcccttgaa gg</pre>	tgccctcgg ggggcaggg cgtgtcctg cggagccgc gaagctgag cagcgggag ccgcgggag ctcactctc ggattctgct gggctccag gaagatgtgg taccaggcc cgcctgtca attccctca	cctaccccta gcctcagctg ctgtgccgca atgctgagca tacctttaca ctcagcaccc ctcacccagc gaggccagcc cctcttgttg agagccgtgg tggagccagc gagagctgct ggcactgct gcctcgtgtc	gtctatacaa ggaggaatat tcatagctga ccgaggccct gggagctcta acccacctca taaccctggc tctcggccac agccgtgtct gcccagtgcc aaccaagctc gctgccactg ccttctctgc actgggctgc	agtaccctgt tgctgcccgg ggctccccaa ccgtctgtgg cccagtgctg acccctgtcc agccggcagt cccttcctta aaggcagacc cgttgcctgc atgcccggag ctgagtcagc aacccgctgt tcgggaggct	gctactgcca ctgttgagca gaactggcct gctgtggctg atgcgggcct atgcagcgga acccctgctg gtcacttgga ttgaagttgc ctgttgttcc gattggctcc ccacactggg cctgtgtgcc gccccgtct	120 180 240 300 360 420 480 540 600 660 720 780 840 900
<pre><211> 2815 <212> DNA <213> Homo sa <400> 34 gggtcctgga gg ggtctcctgt gg tgaaactact tg cttgatct cc tgccccaga gg cctcctatgg tg tagcctcact gg tagcctcact gg tagcctcact gg tagcctcact gg cacaggtgt tg tagcctcact gg cacaggtgt tg cacaggtgt tg cacaggtgt tg cacaggtgt tg cacaggtgt tg tgccagac tg tgccagac tg tgccagac tg tgccagac tg tgcagacttgag gg cagccttgaa gg cagccttgaa gg cagtctggct gg</pre>	tgccctcgg ggggcaggg cgtgtcctg cggagccgc gaagctgag ccgcgggag ctcactctc gattctgct ggagatgtgg taccag gagatgtgg taccag ccgctgtca attccctta ctccccca gctcacct	cctacccta gcctcagctg ctgtgccgca atgctgagca tacctttaca ctcagcaccc ctcacccagc gaggccagcc cctcttgttg agagccgtgg tggagccagc gagagctgct ggcactgct gcctcgtgtc cacccttccc	gtctatacaa ggaggaatat tcatagctga ccgaggccct gggagctcta acccacctca taaccctggc tctcggccac agccgtgtct gcccagtgcc aaccaagctc gctgccactg ccttctctgc actgggctgc attcctcact	agtaccctgt tgctgcccgg ggctccccaa ccgtctgtgg cccagtgctg acccctgtcc agccggcagt cccttcctta aaggcagacc cgttgcctgc atgcccggag ctgagtcagc acccgctgt tcgggaggct gccctcctct	gctactgcca ctgttgagca gaactggcct gctgtggctg atgcgggcct atgcagcgga acccctgctg gtcacttgga ttgaagttgc ctgttgttcc gattggctcc ccacactggg cctgtgtgcc gccccgtct ctcttcttaa	120 180 240 300 360 420 480 540 600 660 720 780 840 900 960
<pre><211> 2815 <212> DNA <213> Homo sa <400> 34 gggtcctgga gg ggtctcctgt gg tgaaactact tg cttgatct cc tgccccaga gg cctcctatgg cc tgcaggtggt gg tagcctcact gg aaaccatcag tg cacaggtgt tg tgtcagacc tg tgcaggtgt tg cacaggtgt tg cacaggtgt tg tgtcagacc tg tgcagacctggga gg cagccttgaa gg cagccttgaa gg cagtctggc gg taccctggc gg taccctggc gg</pre>	tgccctcgg ggggcaggg cgtgtcctg cggagccgc gaagctgag ccgcgggag ctcactctc gattctgct ggagttcag taccag taccaggcc cgcctgtca attccccca gctcaccta	cctacccta gcctcagctg ctgtgccgca atgctgagca tacctttaca ctcagcaccc ctcacccagc gaggccagcc cctcttgttg agagccgtgg tggagccagc gagagctgct ggcactgct gcctcgtgtc cacccttccc aggggctgtg	gtctatacaa ggaggaatat tcatagctga ccgaggccct gggagctcta acccacctca taaccctggc tctcggccac agccgtgtct gcccagtgcc aaccaagctc gctgccactg ccttctctgc actgggctgc attcctcact tggccagctg	agtaccctgt tgctgcccgg ggctccccaa ccgtctgtgg cccagtgctg acccctgtcc agccggcagt cccttcctta aaggcagacc cgttgcctgc atgcccggag ctgagtcagc aacccgctgt tcgggaggct gccctcctct gctgccatat	gctactgcca ctgttgagca gaactggcct gctgtggctg atgcgggcct atgcagcgga acccctgctg gtcacttgga ttgaagttgc ctgttgtcc gattggctcc ccacactggg cctgtgtgcc gccccgtct ctcttcttaa tggctgccc	120 180 240 300 360 420 480 540 600 660 720 780 840 900 960 1020
<pre><211> 2815 <212> DNA <213> Homo sa <400> 34 gggtcctgga gg ggtctcctgt gg tgaaactact tg cttgatct cc tgccccaga gg cctcctatgg cc tgcaggtggt gg tagcctcact gg aaaccatcag tg cacaggtgt tg tgtcagacc tg tgtcagacc tg tgcagacctggg gg agccttgaa gg cagcttgaa gg cagtctggc gg ggactccag agg ggactccag agg</pre>	tgccctcgg ggggcaggg cgtgtcctg cggagccgc gaagctgag ccagggggg ctcactctc gggctcccag gagatgtgg taccaggcc cgcctgtca attccctta ctccccca gctcagcct agatcaca	cctaccccta gcctcagctg ctgtgccgca atgctgagca tacctttaca ctcagcaccc ctcacccagc gaggccagcc cctcttgttg agagccgtgg tggagccagc gagagctgct ggcactgct gcctcgtgtc cacccttccc aggggctgtg tcagggctgtg	gtctatacaa ggaggaatat tcatagctga ccgaggccct gggagctcta acccacctca taaccctggc tctcggccac agccgtgtct gcccagtgcc aaccaagctc gctgccactg ccttctctgc actgggctgc attcctcact tggccagctg ggctcctggg	agtaccctgt tgctgcccgg ggctccccaa ccgtctgtgg cccagtgctg acccctgtcc agccggcagt cccttcctta aaggcagacc cgttgcctgc atgcccggag ctgagtcagc acccgctgt tcgggaggct gccctcctct gctgccatat gctgcccac	gctactgcca ctgttgagca gaactggcct gctgtggctg atgcgggcct atgcagcgga acccctgctg gtcacttgga ttgaagttgc ctgttgtcc gattggccc ccacactggg cctgtgtgcc gccccgtct ctcttcttaa tggctgccc acctcacacc	120 180 240 300 360 420 480 540 600 660 720 780 840 900 960 1020 1080
<pre><211> 2815 <212> DNA <213> Homo sa <400> 34 gggtcctgga gg ggtctcctgt gg tgaaactact tg cttgatct cc tgccccaga gg cctcctatgg cc tgcaggtggt gg tagcctcact gg aaaccatcag tg cacaggtgt tg tgtcagacc tg tgcaggtgt tg cacaggtgt tg cacaggtgt tg tgtcagacc tg tgcagacctggga gg cagccttgaa gg cagccttgaa gg cagtctggc gg taccctggc gg taccctggc gg</pre>	tgccctcgg ggggcaggg cgtgtcctg cggagccgc gaagctgag ccgcgggag ctcactctc gggctccag gagatgtgg taccaggc cagctgtca attccctta ctccccca gctcaccta attccctta ggtcagcct	cctacccta gcctcagctg ctgtgccgca atgctgagca tacctttaca ctcagcaccc ctcacccagc gaggccagcc cctcttgttg agagccgtgg tggagccagc gagagctgct ggcactgct gcctcgtgtc cacccttccc aggggctgtg tccagtgtg tccagtgtg	gtctatacaa ggaggaatat tcatagctga ccgaggccct gggagctcta acccacctca taaccctggc tctcggccac agccgtgtct gcccagtgcc aaccaagctc gctgccactg ccttctctgc actgggctgc attcctcact tggccagctg ggctcctggg ccacctgcag	agtaccctgt tgctgcccgg ggctccccaa ccgtctgtgg cccagtgctg acccctgtcc agccggcagt cccttcctta aaggcagacc cgttgcctgc atgcccggag ctgagtcagc acccgctgt tcgggaggct gcctcctct gctgccaca tacctggcac	gctactgcca ctgttgagca gaactggcct gctgtggctg atgcgggcct atgcagcgga acccctgctg gtcacttgga ttgaagttgc ctgttgtcc gattggccc ccacactggg cctgtgtgcc gccccgtct ctcttcttaa tggctgccc acctcacacc tcgctctggc	120 180 240 300 360 420 480 540 600 660 720 780 840 900 960 1020

ggccttggcc	ctgctgagcc	ggctgctgcc	cggaagtgag	tacctcaccc	atgagctgct	1260
gctgagctgt	gtattccggc	tggagttcct	cccggaaaga	acatcagggg	gtccagaggc	1320
agccgacttc	tctgaccagc	tgtcgttagg	aagcagcaga	gtccctcggt	gtgggcaagg	1380
gactctgctg	gctcaggcct	gccaggacct	ccccagcatc	cgcaactgct	acctgactca	1440
ttgctcgcca	gcccgagcca	gtctgctggc	ctcccaggct	ctgcaccgag	gggagctaca	1500
gcgagtccca	accctgctac	tgcccatgcc	tacggagccg	ctgctgccca	ccgactggcc	1560
cttcctgcac	tgattcgcct	ctacaccggg	cttcagacac	cccctcggga	ctctctccac	1620
agacaccatg	ggcacagcca	tgcgggtcct	gcagtgggtg	ctagttttgg	agagctggcg	1680
ccccaggct	ctctgggctg	tgccccctgc	tgcccgcctg	gcacggctca	tgtgtgtgtt	1740
cctggtggac	agtgagctgt	tccgggagtc	cccagtacag	catctggtgg	cagccctcct	1800
cgcccagctc	tgtcagcctc	aagtcttgcc	aaacctcaac	ctggactgcc	gactccctgg	1860
cctgacgtct	ttccctgacc	tctatgccaa	cttcctggat	cattttgagg	ctgtctcttt	1920
tggggaccac	ctctttgggg	ccctggtcct	cctgcccctg	cagcgtcggt	tcagtgtcac	1980
cttgcgcctt	gccctctttg	gggaacacgt	gggagccttg	cgagctctga	gcctgcctct	2040
gacccagttg	cctgtgtccc	tggagtgtta	cacagtgcct	cctgaagaca	acctggccct	2100
				cgcccacgtt		2160
				ttctctcagg		2220
				aaaacatggc		2280
				aattccacgc		2340
ctttgagctc	tattctcagt	tgccccctct	gcgtcagcac	tacctccaga	gactgacttc	2400
				tgatatagat		2460
				cagacgaaga		2520
				tcctgggagc		2580
agctgtttta	gagcagtgag	ccctaccatt	acatcctgat	atctggggct	tctgaaggtc	2640
tgtgctggga	gtgaagagtg	gcttagctat	ttacccgctc	tttggggaca	gggcaaacta	2700
aatgcatccc	ttcttaccta	actcccaacc	cctgccctgg	gctgaggcat	atgaatgcta	2760
tagttgtgga	*****					2815
Lagitytyta	llaaaalaaa	igillillat	ctcctggaaa	aaaaaaaaa	aaaaa	2013
tagitgigea	Claadatada	igitititat	ctcctggaaa	aaaaaaaaa	aaaaa	2013
cagicgigea	Cladadida	tgtttttat	ctcctggaaa	aaaaaaaaa	aaaaa	2815
<210> 35	ttdadatada	tgtttttat	ctcctggaaa	aaaaaaaaa	aaaa	2813
	llaadalada	tgtttttat	ctcctggaaa	aaaaaaaaa	aaaaa	2613
<210> 35	ttaaaataaa	tgtttttat	ctcctggaaa	aaaaaaaaa	aaaaa	2013
<210> 35 <211> 1078 <212> DNA		tgtttttat	Ctcctggaaa	aaaaaaaaa	aaaaa	2013
<210> 35 <211> 1078		tgtttttat	Ctcctggaaa	aaaaaaaaa	aaaaa	2613
<210> 35 <211> 1078 <212> DNA		tgtttttat	Ctcctggaaa	aaaaaaaaa	aaaaa	2013
<210> 35 <211> 1078 <212> DNA <213> Homo <400> 35	sapiens			ctgcggctgg		60
<210> 35 <211> 1078 <212> DNA <213> Homo <400> 35 ggtgggctct	sapiens gtgctgggtg	ccttcctcac	cttcccaggc		cccagaccca	
<210> 35 <211> 1078 <212> DNA <213> Homo <400> 35 ggtgggctct ccgggacgca	sapiens gtgctgggtg ctgaccatgt	ccttcctcac cggaggacag	cttcccaggc acccatgctg	ctgcggctgg	cccagaccca tgcacaccag	60
<210> 35 <211> 1078 <212> DNA <213> Homo <400> 35 ggtgggctct ccgggacgca cttcctgtct	sapiens gtgctgggtg ctgaccatgt cccctgttca	ccttcctcac cggaggacag tcctgtggct	cttcccaggc acccatgctg ctggacaaag	ctgcggctgg cagttcctcc	cccagaccca tgcacaccag gggacttcct	60 120
<210> 35 <211> 1078 <212> DNA <213> Homo <400> 35 ggtgggctct ccgggacgca cttcctgtct gcaccagccg	sapiens gtgctgggtg ctgaccatgt cccctgttca ccgtttgggg	ccttcctcac cggaggacag tcctgtggct agacgcgttt	cttcccaggc acccatgctg ctggacaaag ctccctgctg	ctgcggctgg cagttcctcc cccattgcac tccgattctg	cccagaccca tgcacaccag gggacttcct	60 120 180 240
<210> 35 <211> 1078 <212> DNA <213> Homo <400> 35 ggtgggctct ccgggacgca cttcctgtct gcaccagccg tgggcgcctc	sapiens gtgctgggtg ctgaccatgt cccctgttca ccgtttgggg tggttgctgg	ccttcctcac cggaggacag tcctgtggct agacgcgttt tggtgctgtg	cttcccaggc acccatgctg ctggacaaag ctccctgctg cctgctgcgg	ctgcggctgg cagttcctcc cccattgcac tccgattctg	cccagaccca tgcacaccag gggacttcct ccttcgactc cccggcccca	60 120 180 240
<210> 35 <211> 1078 <212> DNA <213> Homo <400> 35 ggtgggctct ccgggacgca cttcctgtct gcaccagccg tgggcgcctc cctgcaggccc	sapiens gtgctgggtg ctgaccatgt cccctgttca ccgtttgggg tggttgctgg tacctgtgcc	ccttcctcac cggaggacag tcctgtggct agacgcgttt tggtgctgtg tggccaaggc	cttcccaggc acccatgctg ctggacaaag ctccctgctg cctgctgcgg ccgggtggag	ctgcggctgg cagttcctcc cccattgcac tccgattctg ctggcggtga	cccagaccca tgcacaccag gggacttcct ccttcgactc cccggcccca gggaggctgg	60 120 180 240 300
<210> 35 <211> 1078 <212> DNA <213> Homo <400> 35 ggtgggctct ccgggacgca cttcctgtct gcaccagccg tgggcgcctc cctgcaggcc cctgcaggcc cctgcaggac	sapiens gtgctgggtg ctgaccatgt cccctgttca ccgtttgggg tggttgctgg tacctgtgcc gcccgtgaaa	ccttcctcac cggaggacag tcctgtggct agacgcgttt tggtgctgtg tggccaaggc tccagcagag	cttcccaggc acccatgctg ctggacaaag ctccctgctg cctgctgcgg ccgggtggag ggtggtccga	ctgcggctgg cagttcctcc cccattgcac tccgattctg ctggcggtga cagctgcgaa	cccagaccca tgcacaccag gggacttcct ccttcgactc cccggcccca gggaggctgg atgtgaccgt	60 120 180 240 300 360
<210> 35 <211> 1078 <212> DNA <213> Homo <400> 35 ggtgggctct ccgggacgca cttcctgtct gcaccagccg tgggcgctc cctgcaggcc cctgcaggcc ccgcatcgaa ggtgagcttg	sapiens gtgctgggtg ctgaccatgt cccctgttca ccgtttgggg tggttgctgg tacctgtgcc gcccgtgaaa cagtacctga	ccttcctcac cggaggacag tcctgtggct agacgcgttt tggtgctgtg tggccaaggc tccagcagag cgccgctcat	cttcccaggc acccatgctg ctggacaaag ctccctgctg cctgctgcgg ccgggtggag ggtggtccga cctcaccctc	ctgcggctgg cagttcctcc cccattgcac tccgattctg ctggcggtga cagctgcgaa gtctactgct	cccagaccca tgcacaccag gggacttcct ccttcgactc cccggcccca gggaggctgg atgtgaccgt ttctgctcaa	60 120 180 240 300 360 420
<210> 35 <211> 1078 <212> DNA <213> Homo <400> 35 ggtgggetet ccgggaegea ctteetget gcaccagecg tgggegetet cctgcaggec cctgcaggec cctgcaggec ccgcatcgaa ggtgagettg gaegetggga	sapiens gtgctgggtg ctgaccatgt cccctgttca ccgtttgggg tggttgctgg tacctgtgc gcccgtgaaa cagtacctga ggctattcct	ccttcctcac cggaggacag tcctgtggct agacgcgttt tggtgctgtg tggccaaggc tccagcagag cgccgctcat ggggcytggg	cttcccaggc acccatgctg ctggacaaag ctccctgctg cctgctgcgg ccgggtggag ggtggtccga cctcaccctc	ctgcggctgg cagttcctcc cccattgcac tccgattctg ctggcggtga cagctgcgaa gtctactgct aactgcacac	cccagaccca tgcacaccag gggacttcct ccttcgactc cccggcccca gggaggctgg atgtgaccgt ttctgctcaa cccgacccat	60 120 180 240 300 360 420 480
<210> 35 <211> 1078 <212> DNA <213> Homo <400> 35 ggtgggetet ccgggaegea ctteetget gcaccageeg tgggegetet cctgcagee cctgcagee cctgcagee ccgcatcgaa ggtgagettg gaegetggga cctcagecag cctcagecag	sapiens gtgctgggtg ctgaccatgt cccctgttca ccgtttgggg tggttgctgg tacctgtgc gcccgtgaaa cagtacctga ggctattcct cgctgcccc	ccttcctcac cggaggacag tcctgtggct agacgcgttt tggtgctgtg tggccaaggc tccagcagag cgccgctcat ggggcytggg atcggctctg	cttcccaggc acccatgctg ctggacaaag ctccctgctg cctgctgcgg ccgggtggag ggtggtccga cctcaccctc cccagctcct	ctgcggctgg cagttcctcc cccattgcac tccgattctg ctggcggtga cagctgcgaa gtctactgct aactgcacac ctactatccc	cccagaccca tgcacaccag gggacttcct ccttcgactc cccggcccca gggaggctgg atgtgaccgt ttctgctcaa cccgacccat actgcagcgc	60 120 180 240 300 360 420 480 540
<210> 35 <211> 1078 <212> DNA <213> Homo <400> 35 ggtgggetet ccgggaegea ctteetget gcaccageeg tgggegete cctgeagec cctgcagec ccgcatcgaa ggtgagettg gacgetggga cctcagecag ggattgeegg	sapiens gtgctgggtg ctgaccatgt cccctgttca ccgtttgggg tggttgctgg tacctgtgc gcccgtgaaa cagtacctga ggctattcct cgctgcccc	ccttcctcac cggaggacag tcctgtggct agacgcgttt tggtgctgtg tggccaaggc tccagcagag cgccgctcat ggggcytggg atcggctctg	cttcccaggc acccatgctg ctggacaaag ctccctgctg cctgctgcgg ccgggtggag ggtggtccga cctcaccctc cccagctcct gggaggacga ctcccctct	ctgcggctgg cagttcetcc cccattgcac tccgattctg ctggcggtga cagctgcgaa gtctactgct aactgcacac ctactatccc agtccagcag	cccagaccca tgcacaccag gggacttcct ccttcgactc cccggcccca gggaggctgg atgtgaccgt ttctgctcaa cccgacccat actgcagcgc	60 120 180 240 300 360 420 480 540
<210> 35 <211> 1078 <212> DNA <213> Homo <400> 35 ggtgggetet ccgggaegea ctteetget gcaccageeg tgggegetet cctgcaggec cctgcaggec cctgcategaa ggtgagettg gaegetggga ccteagecag ggattgeegg acctertetg	sapiens gtgctgggtg ctgaccatgt cccctgttca ccgtttgggg tggttgctgg tacctgtgcc gcccgtgaaa cagtacctga ggctattcct cgctgcccc ggcyctgggt gtggacggct	ccttcctcac cggaggacag tcctgtggct agacgcgttt tggtgctgtg tggccaaggc tccagcagag cgccgctcat ggggcytggg atcggctctg ggcctgctta gcctgccagc	cttcccaggc acccatgctg ctggacaaag ctccctgctg cctgctgcgg ccgggtggag ggtggtccga cctcaccctc cccagctcct gggaggacga ctcccctctt	ctgcggctgg cagttcctcc cccattgcac tccgattctg ctggcggtga cagctgcgaa gtctactgct aactgcacac ctactatccc agtccagcag cctccgtggc	cccagaccca tgcacaccag gggacttcct ccttcgactc cccggcccca gggaggctgg atgtgaccgt ttctgctcaa cccgacccat actgcagcgc gtcctggcct ctctacttcc	60 120 180 240 300 360 420 480 540 600 660
<210> 35 <211> 1078 <212> DNA <213> Homo <400> 35 ggtgggetet ccgggaegea ctteetget gcaccagceg tgggegete cctgeagge cctgeategaa ggtgagettg gacgettgga cctcagcag gattgegg acctcatett accagcag	sapiens gtgctgggtg ctgaccatgt cccctgttca ccgtttgggg tggttgctgg tacctgtgcc gcccgtgaaa cagtacctga ggctattcct cgctgcccc ggcyctgggt gtggacggct ggcaggctcc	ccttcctcac cggaggacag tcctgtggct agacgcgttt tggtgctgtg tggccaaggc tccagcagag cgccgctcat ggggcytggg atcggctctg ggcctgctta gcctgccagc tagctgcctg	cttcccagge acccatgctg ctggacaaag ctccctgctg cctgctgcgg ccgggtggag ggtggtccga cctcaccctc cccagctcct gggaggacga ctcccctctt tgctcgccag cagaccctcc	ctgcggctgg cagttcctcc cccattgcac tccgattctg ctggcggtga cagctgcgaa gtctactgct aactgcacac ctactatccc agtccagcag cctccgtggc	cccagaccca tgcacaccag gggacttcct ccttcgactc ccttcgactc gggaggctgg atgtgaccgt ttctgctcaa cccgacccat actgcagcgc gtcctggcct ctctacttcc aggtctgttc	60 120 180 240 300 360 420 480 540 600 660 720
<210> 35 <211> 1078 <212> DNA <213> Homo <400> 35 ggtgggetet cegggaegea ctteetget geaceageeg tgggegete cetgeagee cetgeagee cegeategaa ggtgagettg gaegetgga ceteageag gattgegg acetertetg aceageagt tegggegete cetgaggae ceteageag	sapiens gtgctgggtg ctgaccatgt cccctgttca ccgtttgggg tggttgctgg tacctgtgcc gcccgtgaaa cagtacctga ggctattcct cgctgcccc ggcyctgggt gtggacggct ggcaggctcc gggacactag	ccttcctcac cggaggacag tcctgtggct agacgcgttt tggtgctgtg tggccaaggc tccagcagag cgccgctcat ggggcytggg atcggctctg ggcctgctta gcctgccagc tagctgcctg	cttcccagge acccatgctg ctggacaaag ctccctgctg cctgctgcgg ccgggtggag ggtggtccga cctcaccctc cccagctcct gggaggacga ctcccctctt tgctcgccag cagaccctcc	ctgcggctgg cagttcctcc cccattgcac tccgattctg ctggcggtga cagctgcgaa gtctactgct aactgcacac ctactatccc agtccagcag cctccgtggc ccttttcggc	cccagaccca tgcacaccag gggacttcct ccttcgactc cccggcccca gggaggctgg atgtgaccgt ttctgctcaa cccgacccat actgcagcgc gtcctggcct ctctacttcc aggtctgtca	60 120 180 240 300 360 420 480 540 600 660 720 780
<210> 35 <211> 1078 <212> DNA <213> Homo <400> 35 ggtgggetet ccgggaegea ctteetget gcaccagceg tgggegete cctgcaggec cctgcaggec ccgcatcgaa ggtgagettg gacyctggga cctcagcag ggattgccgg acctertett accagcact ctggggeact ctggggeace ggtgggeace ggtgggeace ggtggggecg	sapiens gtgctgggtg ctgaccatgt cccctgttca ccgtttgggg tggttgctgg tacctgtgcc gcccgtgaaa cagtacctga ggctattcct cgctgcccc ggcyctgggt gtggacggct ggcaggctcc gggacactag gactcccgg	ccttcctcac cggaggacag tcctgtggct agacgcgttt tggtgctgtg tggccaaggc tccagcagag cgccgctcat ggggcytggg atcggctctg ggcctgctta gcctgccagc tagctgcctg cctgccect	cttcccaggc acccatgctg ctggacaaag ctccctgctg cctgctgcgg ccgggtggag ggtggtccga cctcaccctc cccagctcct gggaggacga ctcccctctt tgctcgcag cagaccctcc ctgtttgcgc accacagtgc	ctgeggetgg cagtteetee cecattgeae teegattetg ctggeggtga cagetgegaa gtetactget aactgeaeae ctactateee agteeageag ceteegtgge cettttegge tggggeeetg cecegtgtee ctgaeeegg	cccagaccca tgcacaccag gggacttcct ccttcgactc cccggcccca gggaggctgg atgtgaccgt ttctgctcaa cccgacccat actgcagcgc gtcctggcct ctctacttcc aggtctgttc ccagctgcaa gcccccttg	60 120 180 240 300 360 420 480 540 600 660 720 780 840
<210> 35 <211> 1078 <212> DNA <213> Homo <400> 35 ggtgggctct ccgggacgca cttcctgtct gcaccagccg tgggcgctc cctgcaggc cctgcaggct ccgcatcgaa ggtgagcttg gacgctgga cctcagccag gattgccg acctcrtctg accagcact tctggggcacc ggtgggccg ggtgggccg ggtgggccg gacgccgagt	sapiens gtgctgggtg ctgaccatgt cccctgttca ccgtttgggg tggttgctgg tacctgtgcc gcccgtgaaa cagtacctga ggctattcct cgctgcccc ggcyctgggt gtggacggct ggcaggctc gggacactag gactcccgg	ccttcctcac cggaggacag tcctgtggct agacgcgttt tggtgctgtg tggccaaggc tccagcagag cgccgctcat ggggcytggg atcggctctg ggctgctta gcctgccagc tagctgccagc tagctgcctg cctgccact cgttcccttc	cttcccagge acccatgctg ctggacaaag ctccctgctg cctgctgcgg ccgggtggag ggtggtccga cctcaccctc cccagctcct gggaggacga ctcccctctt tgctcgccag cagacectcc ctgtttgcgc accacagtgc tcctgggecc	ctgcggctgg cagttcctcc cccattgcac tccgattctg ctggcggtga cagctgcgaa gtctactgct aactgcacac ctactatccc agtccagcag cctccgtggc ccttttcggc tggggccctg ccccgtgtcc	cccagaccca tgcacaccag gggactcct ccttcgactc cctcggcccca gggaggctgg atgtgaccgt ttctgctcaa cccgacccat actgcagcgc gtcctggcct ctctacttcc aggtctgttc ccagctgcaa gcccccttg	60 120 180 240 300 360 420 480 540 600 660 720 780 840 900
<210> 35 <211> 1078 <212> DNA <213> Homo <400> 35 ggtgggetet ccgggaegea ctteetget gcaccagceg tgggegete cctgeagge cctgeategaa ggtgagettg gacgetgga cctcagcag gattgegg acctertetg accagcat ggtgggeet tctggggeget ggtgggeeg gacgegage ggtggggeeg gacgeegagt gcattgtet	sapiens gtgctgggtg ctgaccatgt cccctgttca ccgtttgggg tggttgctgg tacctgtgcc gcccgtgaaa cagtacctga ggctattcct cgctgcccc ggcyctgggt gtggacggct ggcaggctcc gggacactag gactccccgg ttctgcctca ccgaagcgta	ccttcctcac cggaggacag tcctgtggct agacgcgttt tggtgctgtg tggccaaggc tccagcagag cgccgctcat ggggcytggg atcggctctg ggcctgctta gcctgccagc tagctgcctg cctgcccct cgttcccttc gaactgtctc tgtgccaggt	cttcccagge acccatgctg ctggacaaag ctccctgctg cctgctgcgg ccgggtggag ggtggtccga cctcaccctc cccagctcct gggaggacga ctcccctctt tgctcgccag cagaccctcc ctgtttgcgc accacagtgc tcctgggccc ttgagtggcr	ctgeggetgg cagtteetee cccattgeae tccgattetg ctggeggtga cagetgegaa gtetaetget aactgeaeae ctactateee agteeageag ceteegtgge cettttegge tggggeeetg cecegtgtee ctgaeeegeg ageageatga	cccagaccca tgcacaccag gggacttcct ccttcgactc cctcggccca gggaggctgg atgtgaccgt ttctgctcaa cccgacccat actgcagcg gtcctggct ctctacttcc aggtctgttc ccagctgcaa gcccccttg gggtcccgag tggctgctct	60 120 180 240 300 360 420 480 540 600 660 720 780 840 900 960

```
<210> 36
<211> 1217
<212> DNA
<213> Homo sapiens
<400> 36
cggcacgagg ttgaatgtta gccctggagg agatccatgt cttactcgct ctttctggcc
                                                                      60
cttctgtctt ttgcctctgc aattcttttt gtagctggca cgatagcagg gactgggggt
                                                                     120
ctatcctttc atggtattgc tacaatattt gtccttactg gaaaatggta acatccgggt
                                                                     180
ctgatttaat tggcattaca cttacacagg gactctgagc acccccgtca ccacaccaga
                                                                     240
cagtggacca gttttcacag ctacaaagag ctagaaatgt gtttaacatc atccagtgca
                                                                     300
tcccctaatt caaaaccatc ctcactaatc aatcatattc acccataaat attacaaatg
                                                                     360
agattgattc catctcaaga caatttgtca aatacttaat tttcttcctg gatgattcta
cttactggat attttagaaa gagaaatgtc tgagataaaa tccctcacat ttactcaata
                                                                     480
taacaaatta ctgtttctac tcctattctg agtagtgctt ctgaagattg tttgctgtag
tgttgtcttt gataaaatga atgtcagtag tgagcctttt agagatacca tgctcagaaa
                                                                     600
tectetttgg gateagaaga tacetaaaat teteceettt tgeecaettg gttagatgag
                                                                     660
720
agtggttgta tctgtggctg tgatggttgt tgttacttgt ctctctct ctctggctct
                                                                     780
ggcttttgct ttcctgctag tgttctttct ctttccaaac aaatagttaa attaaacgtg
                                                                     840
agcttctgaa ttgtacttgt tcatactttc aaaacataac agattaataa aaatagatgt
                                                                     900
gtcctgattt aaaacatgcc ccctggaaag gcatgctgta ttatgaaatc atgataatat
                                                                     960
aactgcatta ttacatggca gtataaatat tagtctgttg aattcatttg tccaattgta
                                                                    1020
taactttgtg gagcagtgtt ttgacctttg atacataatt ctggagcaag tggagtggtt
                                                                    1080
gcaggcagat gagacagtgt tatatcagga tttttcaatc aactttagtt ggaggcctgg
                                                                    1140
caattacaaa catcttcaga tgtttctgta accattataa atatgaaaaa aacctcttca
                                                                    1200
aaaaaaaaa aaaaaaa
                                                                    1217
<210> 37
<211> 1282
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (153)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (1220)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (1222)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (1232)
<223> n equals a,t,g, or c
<220>
```

```
<221> SITE
<222> (1246)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (1282)
<223> n equals a,t,g, or c
<400> 37
actogtgccg aattoggcac gagccattot gagtttggto cottoccaaa agtaggggtt
                                                                       60
ttqtgtggaa aatctgagca aacctctgtt gactgttctg gggtggagtg aagggagaar
                                                                      120
                                                                      180
gggctcagct aaagaacatg gggagattag ggnaacaatg ccttttattt cttgctttta
                                                                      240
aagcaatttc aggagttttc ttcctctttt ggcgtcctgc tgactccaca gagcggaaca
                                                                      300
cccaaagctg ggactttcca cctctctaat gctcagtgaa gagcgggcca ggggggtgtg
                                                                      360
gaaaagaaag ggtcctggag gagcccaaat tacgaatggc tagagactgg cattggcaag
                                                                      420
cgaggaggct tcgtcacagt gtagtcttcc ggttgtccga gggtactgtc ccaggggctg
                                                                      480
gggggtrttc cgtcttctgc agatcaactc ccgcaggcta aatgtggaca tcgcggtatc
                                                                      540
atgcttgata aacggaccaa taatcaagtg gagattcatt agaaccacat aacccatact
                                                                      600
aggitgatti cicaagtata agscctggic tgttgcccag sctggagtgc actgacacca
                                                                      660
tcatggttca ctgcagcctc aaactcctga gcccaagtga tycctcccac tcagcctcac
aagtagctaa gactagaggt gtgcaccatc amacccagct aatttttaaa gttttttttg
                                                                     720
                                                                     780
tararatggg gtcccactct acaaaatatw taagtataag gcckggtctg ttgcccargc
                                                                      840
tgggsaaccc ttggactaag gcaatcctcc agcctcagcc tcctaaagtg ctgggattac
                                                                     900
aggcgtgagc caccgcaccc acctctagga tctctactat tgaggaaaaa ttggaggcat
                                                                     960
caaactccaa gggcaaaaca tgaagactcg ctggcccacc atggatggag gttttctctc
                                                                     1020
ttaaaattcc cacagcaccg catggaactg cctctcctgg gacctcagcg tttccttctt
                                                                    1080
tgctctaagc aatagcctct gccactggag attctgagat ggccgatttc cttttggata
                                                                    1140
tttaagtttt gaaatcatgc tcatttggca taggaatgtt tcacttcagt ctcctttaaa
                                                                    1200
caaaaggaca cacaaccacg attgcccctc cctcccgaag ggtcactgga cttcatgcat
                                                                    1260
cagtaatgtt tccaaaaatn tnttaagtac cnacatgcag tggccngctt ttcatttttc
                                                                    1282
caagtgaagc catcagaaaa an
<210> 38
<211> 559
<212> DNA
<213> Homo sapiens
<400> 38
gatteggeac gagetgaage eetgggtgee actgetggee cageagggag gaggttgetg
                                                                       60
                                                                     120
ctgctcgggc tgaagtgagg tgtgggtctg gctgggcctc cagtttccca cctgggcctt
gattgtgagg aaggcctggc ctggctgcag aagcccagaa gcacctgagt aggagagttc
                                                                     180
                                                                      240
ctttgtccca cctgcagctc attcaagcct gtgcatgggg gttggggtcc tcaggatctt
gctttcctgt ttaggggagg cagccccaaa gagtgctggg accagtttgg agagtgctaa
                                                                     300
                                                                     360
ggaatgctgg tctgcagcga ccctacttgt gctctgcgtc ctctgccaac tgcagcatgg
                                                                     420
gtgaacatct gtacatctgt ccccataatg aaaatggcct cagcaaataa caaaaatatt
480
                                                                     540
aaaaactcga gggggggccc ggtacccaat tcgccctata gtgagtcgta ttacgcgsgs
                                                                     559
tcamtggccg tcgtttaca
```

<210> 39

<211> 803

<212> DNA

```
<213> Homo sapiens
<400> 39
                                                                        60
ggcagagcta ggccaggcag agcctagctc ttgccagggc agcaggaagc cacacagtgt
gttgaagccg gagcaggaga gggggccctg actcccatgt gtccttgcag gcaggagcag
                                                                       120
                                                                       180
ttcgtggact tgtacaagga gtttgagcca agcctggtca acagcaccgt ctacatcatg
gccatggcca tccagatggc acctttcgcc atcaattaca aagtaaggcc tgggccctgc
                                                                       240
cmaaacattc actgtctgcc cacccagccc caccccatga agccatctgt ccctcatccc
                                                                       300
cacagggeec geeetteatg gagageetge eegagaacaa geeeetggtg tggagtetgg
                                                                       360
                                                                       420
cagtttcact cctggccatc attggcctgc tcctcggctc ctcgcccgac ttcaacagcc
                                                                       480
agtttggcct cgtggacatc cctgtggagt tcaagctggt cattgcccag gtcctgctcc
                                                                       540
tggacttctg cctggcgctc ctggccgacc gcgtcctgca gttcttcctg gggaccccga
agctgaaagt gccttcctga gatggcagtg ctggtaccca ctgcccaccc tggctgccgc
                                                                       600
tgggcgggaa ccccaacagg gccccgggag ggaaccctgc ccccaacccc ccacagcaag
                                                                       660
gctgtacagt ctcgcccttg gaagactgag ctgggacccc cacagccatc cgctggcttg
                                                                       720
gccagcagaa ccagccccaa gccagcacct ttggtaaata aagcagcatc tgagatttta
                                                                       780
aaaaaaaaa aaaaaaactc gag
                                                                       803
<210> 40
<211> 1510
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (426)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (454)
<223> n equals a,t,g, or c
<400> 40
cacgagaaac attctatctt ttatcaaatg tgtgattcat aacttttgga taccaaagga
                                                                        60
                                                                       120
atctaacgaa ataaccataa tcatcaatcc atacagggag actgtgtgct tctctgtgga
gcctgtcaag aagatattta actatatgat acatgtgaat cgaaacatca tggatttcaa
                                                                       180
actetteett gtgtttgtgg caggagtttt tettttett tatgcaagga ceetggagte
                                                                       240
aaagccctac tttctattac tcctcgggaa ctgtgctagg tgttctaatg acatagtctt
                                                                       300
tgtcttgctg ttggtgaaaa gattcatccg aagtatagca ccttttgggg ctctaatggt
                                                                       360
                                                                       420
tggttgttgg tttgcctcag tttatattgt atgccagttg atggaagatc tgaagtggct
                                                                       480
gtggtntgaa aacaggatat atgtatcagg ctangtcttg atagttggat ttttcagctt
tgttgtttgt tacaagcatg ggccccttgc acacgacagg agcagaagtc ttctgatgtg
                                                                       540
                                                                       600
gatgctgcga ctcctctcc tggttctggt ctatgctggt gtggctgtgc ctcagtttgc
ctatgcagcc ataatcctcc tcatgtcctc ctggagtctg cactacccac tgagagcatg
                                                                       660
                                                                       720
cagttatatg aggtggaaaa tggagcagtg gtttacatca aaagagctgg tggtgaaata
                                                                       780
tcttacggaa gacgagtaca gggagcaagc tgatgctgaa acgaacagtg ctctggagga
                                                                       840
gctacgccgg gcctgccgaa aacccgactt tccctcatgg ctggtcgtct ccagactcca
                                                                       900
cactcctagc aaatttgcag attttgttct tggaggaagc cacttgtcac ctgaagaaat
                                                                       960
cagtctgcat gaagagcagt atggccttgg gggtgccttc ttggaagagc agctctttaa
cccgagtact gcctgacatg cgaccttcaa gttgacttca ttctggacaa ggaagtgggc
                                                                      1020
aaagggcagg attctattaa agttaggcag aactgttcta gtgaacggtg gcaaaaacat
                                                                      1080
ttgctgtgga gaaaaacaag tcagtctgga aaggaaaacc aacccatttt gaagataact
                                                                      1140
tagcattctt ggtgacttct gctacttatt gtactgtagg tggataccaa aattctgtga
                                                                      1200
```

PCT/US99/05804

WO 99/47540

```
caqccactac cacttacctt gaatgaaggc tttcattagg aacaggggaa tggcgttgtt
                                                                    1260
cttaaggggc tagtaagcat gaacaggtgc tttgtcgaca ccagggcact aaatctggtc
                                                                    1320
ttaatcccct gaacctgtgt cagaagactc tgcaatactc ttcctatagt tcgtcagtat
                                                                    1380
aagtccttaa agagacctga gacatgctgg accagtgttt tccaaagtac agctcacagg
                                                                    1440
ctactaccaa gtgttggtca ataaaggtat tctgaggtca actaagattg ataaaaaaaa
                                                                    1500
                                                                    1510
aaaaaaaaa
<210> 41
<211> 1095
<212> DNA
<213> Homo sapiens
<400> 41
qcttqqtqgt gctatttqct tcttcaaatt ctcqttattt aaaatatttc tttcttqtac
                                                                      60
cgttgattct gggatcagcc tggatgtgtc aaacactgcc tgccaggctt agagctcagt
                                                                     120
gcatttcttc ccttttattc ctgctgatgg gattgctggc catgaccggt gagaggaatc
                                                                     180
aaggaaccca ttactatgag ttctcaggat tcatcttcaa atctcaaatg atgtggtcaa
                                                                     240
ttaaaccaaa ttaaaaacaa gctcttgtta aaagcaagtt aaaaacaagc tcttgacctt
                                                                     300
qaqaaqaaat gattggtatt aggaaqactg ttgagctgat actgcccttc attcattctc
                                                                     360
taccctggtg cttggataca ggagcaaagt aagaaaataa tcacagcttt attgagggct
                                                                     420
ctatgagcaa ggcttggtga ggatggaaga gaatggagct atcagttgat gagaacctac
                                                                     480
                                                                     540
taggtgttga gctccttaca ttcattgcct atttaaaaact ttctaacaac ttcatgtgta
agcgttgtcc cgatttaaaa aaaaaaatag atgtggaaac tgaacctgga gaaggtgtgt
                                                                     600
                                                                     660
aatttgtcca aggttgcaca ggcaaagggg caaaattcag ctttaaaccc aggactgttt
ccacagctcc aagtyccctt tattcatggg atttgtaaga tggagcccct gccactgtag
                                                                     720
catttataac ttactttgga gaataagatt cctgaaagta cgtttaataa aaaaaaaaga
                                                                     780
                                                                     840
tgtccagcta tgtacggcag ctcacgcctg taatcccagc actttgagag gcaaaggggg
gaggataget tgaggetaag agtttaagae taacetggge aacatggeaa gaceetgtet
                                                                     900
                                                                     960
ctaaaaaaca aaattagcca gttgtggtgg catgcacctg tagtccaagc tactcaggag
gctaaggtga gagggtcgct tgagcccagg agtttgaggc tgcagtgagc catgatggcg
                                                                    1020
1080
                                                                    1095
aaaaaaaac tcgag
<210> 42
<211> 1162
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (340)
<223> n equals a,t,g, or c
<400> 42
                                                                      60
ggcacgagct gattcctaag gaatattcta gccaaatcat gtatctgtgg tttagttttt
ctacagtagg gctgtgcggt tgctgcctgc tttatagggc atgtgggttt atatggtatc
                                                                     120
tgctgttact tgggcacagc agcaccaact cattacagga tggaggggca gaacgcccag
                                                                     180
agcacccctg ggctcacgtg cggtacagct gcaggagaga gctgtccttt tggttttatg
                                                                     240
tttttaatta attctgtttc ctcagattga tgattaaatt tatttttcca gcctgaccaa
                                                                     300
                                                                     360
gaaggcgtca ccataccaga tctggggagt ctctcctcan ctctgataga cacagagagg
aatctgggcc tgcttctcgg attacacgct tcctatttag caatgagcac accgctgtct
                                                                     420
                                                                     480
cctgtcgaga ttgaatgtgc cagtaagaaa atctttactt tttgctaatt agcagatttt
                                                                     540
ttttttttgg aactgtaagt gccattaaga gtgggagagg gccaggcaca gtggttcatg
```

cctgtaatcc cagcactttg	ggaggttgtg	gcacgtggat	tgcttgagat	caagattttg	600
agaccagcct gggcaacatg	gcaaaacccc	atctctacaa	aaaacacaaa	aattagccag	660
gcatgttggc acgtatttgt	agtcccagat	actcaggagg	ctgaggtagg	aggattgctt	720
gagcctggga ggttgaggct	gcagtgagtc	atgatcatac	cactgcactc	cagcctgggt	780
gacagagcaa gactctctct	ttaaaaaagc	aggagatggc	caggcagtgg	ctcatgcctg	840
taatcccagc actttgggag	gctgaggcgg	gtggatcacc	tgaggtcagg	agttcaagac	900
cagcctggcc aatgtggtga	aaccccatgt	ctactaaaaa	tgcaaaaatt	agctgggtgt	960
ggtgacgggt gcctgtaatc	ctaggtactc	gggacgctga	cgtaggagaa	ttgcttgaac	1020
ccaggacaca gaggttgcag	tgagctgaga	tcacgccact	gcactccagc	ctgggtgaca	1080
agagegagae teggteteca	aaaaaaaaa	aggagaggag	gattcaacac	agttgatgat	1140
gacaaaaaa aaaaaaaaa	aa				1162
<210> 43					
<211> 657					
<212> DNA			•		
<213> Homo sapiens					
<220>					
<221> SITE					
<222> (12)					
<223> n equals a,t,g,	or c				
<400> 43					
ccccccggg gntgcaggaa	ttcggcacar	attttacatg	cttttaagtt	aatgttggaa	60
aactaatcac aagcagtttc	taaaccaaaa	aatgacatgt	tgtaaaagga	caataaacgt	120
tgggtcaaaa tggagcctga	gtcctgggcc	ctgtgcctgc	ttcttttcct	gggaacagcc	180
ttgggctacc caccactccc	aaggcattct	tccaaatgtg	aaatcctgga	agtaagattg	240
caccttcttc ctctcctgat	caacatcggt	atgatgtctc	ctgttgcctc	accctttgtc	300
tgcagtatca ctggatagga	ctggtggaaa	gggagcagcc	tgacagagct	ccaaatgtgg	360
agaatatggc atccctccac					420
ccttatcctg accaaagact					480
atacaatctt acttgcaggt			=		540
caaacatccc ctggttggtg	atcacttaca	gttgtgtcca	cctttatttt	atgtactttg	600
attaaaaaaa aaaaactttt	tgttaatata	aaaaaaaaa	aaaaaaaaa	aaaaaaa	657
<210> 44					
<211> 1155					
<212> DNA					
<213> Homo sapiens					
-Day mana Day Land					
<400> 44					
ggcacgagtg gaagtgtaag	cagaaataca	gcgagggctc	aggaaatact	agaataggca	60
acatgetett cetetetget	tctatctgca	catctgcttt	atttctttgc	ctcagcagac	120
tcaccatctc tgctcctcat	cccgcatggt	ggggaaggat	gcccacccac	acctccccag	180
gccatctgtt agagctccaa					240
ctctagtttc aaattcctgg	ggaaaaatga	cccagctcac	ttcaggctcc	cactcttggt	300
ccagtgggct gcaaaatttc					360
aaaatgaagg gaataattgt					420
ggggagaatg cccaggggac					480
ggaaaatccc aattttattt	tcctacagag	tcagcatccc	acacattttc	cttcacagaa	540
actgacaaat aatccatggg					600
catcagtttt caacaccttg					660
agcacaatgc atctccctct	aattgtgtca	tgtgctggag	gagaatgtga	agttctgtct	720

gtctttagca aac ttgatgatcc act aggattctgt gta taggtagctg tag cacttactca gaa	tgatttt actttaag gttaggat	agttttagga gaatgttttc tttttggttg	cctggatgca tgagcatgtc taagtatgtg	taggcagatg cagtacaaca aagatttaaa	tcagtttaca gacgctctgt tgtatcagct	780 840 900 960 1020
cagtatttgg cct cacaagcttt ctc aaaaaaaaaa aaa	tatgtgg					1080 1140 1155
<210> 45 <211> 1112 <212> DNA <213> Homo sar <400> 45	oiens					
gccggaggaa gag	reatetae	aaaactgggt	tcctagaagt	atagacggac	ttagcttttw	60
gtagaatttg gtg						120
teceggeage acg						180
agcctgtccc cta						240
cctctctgtt tto	ttgggcc	taggcaagtt	ttagaggagt	tcccaagcat	tgttatgagg	300
ccagtgtgct cgc	tgggctg	ggcgggatgg	cctgggcttg	tgtgtggcct	gagggctctc	360
ctggggcctt ctc	ttttccc	agtcaccttt	ggagccacag	aagcagtgca	ctcattggat	420
gtctgttctt aac						480
agcagtgctc ato						540
cccgctgaaa ctg						600
ttctccatgt cac						660
atactgtgta gga						720
ccgatgatag tad						780
atgtacagat tat						840
gattttttt tac			•			900
atgaatgttt att	_					960
ttgcatatgt atg						1020
ggttttatac aaa	=			acaatcatgt	aattaaaagt	1080
attaacccaa aaa	aaaaaaa	aaaaaaytcg	ag			1112
<210> 46						
<211> 4023						
<212> DNA						
<213> Homo sar	piens					
<220>						
<221> SITE						
<222> (1049)						
<223> n equals	a,t,g,	or c				
<220>						
<221> SITE						
<222> (2758)						
<223> n equals	a,t,g,	or c				
<400> 46						
cccacgcgtc cgt	ccaaaca	tcaggaggca	ggcagcatgg	taaatgagaa	agaagccagg	60
actgggagtc caa	agtcctg	gcttctatgt	ctggctttgc	tactaatcaa	atatgtgact	120

	_					
	catacctcac					180
	ccccttgga					240
	gtcagtccgc		_			300
	tctctgtgat					360
	tggtgggtct					420
atctatcacc	taccatgtga	ttacagtgca	gattttttt	tccttttcct	tttcttttc	480
tttcttttt	tttttttt	tgtttgagac	ggagtctcgc	tttgtcaccc	aggctgcagc	540
gcagtggcgc	gatctcggct	cactgcaagc	teegeeteee	gggttcaccg	ccattctcct	600
	cccgagtagc					660
tctattttta	gtagagacag	ggtttcaccg	tgttagccag	gatggtctcc	atctcctgac	720
gtggtgatcc	gcccgcctcg	ggctcccaaa	gtgctgggat	tacgggcgtg	agccactgcg	780
	agtgcggata			_		840
caacccttaa	tacatactat	cgtatgaaat	gcctctttcc	aaattcagcc	ttttctaaaa	900
	gaaaactgct					960
gggagaatag	tttttcatgt	ttctttgttt	aagcaataga	aagaaaggaa	acttatgtcg	1020
tttacttttc	aggccataga	ggttttcana	acaacttgaa	aacatgatca	aattagccaa	1080
acttctgata	gttttcaatg	tagtctgtga	tcatgggata	atttagcctc	agttctttt	1140
ctgaaattgt	gttttgaatg	tttgatttga	cttatttacc	atcaaacttg	ctataaggtt	1200
attactctaa	tgaataagca	tattccctta	attgggacaa	tttactatta	tttctttcat	1260
aaagtagggc	accattcacc	atctatttcc	tggctcttta	gttatcaaaa	tgttaagctc	1320
attgctattc	atcccggcac	agcacttata	tgagaggcat	gaagctggct	gaattctgca	1380
tcattaggaa	tgacacagcc	tcatcacatt	gacaccagtg	cttgtctctc	acaccaatcc	1440
	caactgaaaa			-		1500
atgttttta	gtctggtggt	acccaaaatt	gaacaaaaa	tgggtgctgc	ttctcttaat	1560
aggcaaaact	atgctgcagg	ataatgtatt	catgcagggt	cttccagcca	gaccccaaat	1620
catccctccc	ttcactagaa	tttttctgtt	taattcgatg	gccactctcc	acagggatcc	1680
attctgtgtc	ttattacagg	agatgctcaa	tgaatgaggg	acttatcttc	tagaaatgca	1740
gctccgaggt	agtctgttga	gtgaaataat	gaatccatta	tcacaaaata	aattgaaagc	1800
tgtctgacat	ttggacaatt	tttattttgt	ttcacattgt	tctgaaaact	atactgtttc	1860
ttttctccct	attatttaaa	taagcaaatg	atgaacagat	tacaaaattg	aggacactcg	1920
aggtagggaa	ggagcccctc	gacaggagga	tcaggacata	gtaccaaggg	caagagaaac	1980
gattcaataa	acactattta	ctatatattt	taggcatggt	tctaggtaat	cacatgataa	2040
gtagttgaaa	gaactgaaaa	tgttttatct	gcaagaaaag	ggcaagtgta	atatcttcaa	2100
	gaatgtaaat					2160
	ataggctgaa					2220
	ggtgagagtt					2280
	ttataaggtg					2340
	taactctagg					2400
cctgccatca	agaacatgtt	cctcatctgc	aggacaggac	caagatggct	ctgtctacct	2460
	ctgtgaggcg					2520
-	taaatgtgta				-	2580
	ctttcttct					2640
-	gcttttcagg			-		2700
	gtggaggcct					2760
	cagcctgctc					2820
	tettteeet					2880
caaacagtgc	ttttctcaca	gaactactgg	aaagagtcct	ggctcggaaa	acttgctctt	2940
gaaagtggca	cggccagagc	aggggtctct	agagggtcgt	gccacctcta	cctgccacag	3000
	cggtcaggta					3060
gcagcctggg	gctgctcctg	agtcatgttt	ccacttctgt	cttacaggcc	tcattttcct	3120
	ctgtaaaaat					3180
	ttttttttt				-	3240
acgttcatag	gaagcaacaa	aagttgatct	cttggtctct	acaataggag	aggattgggc	3300
	caaagctgac					3360
ctgcccctat	catgtcaggc	agtgagtaca	gttaaaacat	tgggaattgg	taaaggaaag	3420

```
aaaaaaactg aaaagaaccc tttgaagtta gacaaactgt ccagagacat agtgctaaaa
                                                                     3480
tectecytet ttttettyce acagetteta gaatteetet ecagagetae teteaagtta
                                                                     3540
tatccagggg acaggccct ttggctccaa cccacacgcc tgaactttaa ggatcattgg
                                                                     3600
actatettet etgtggeear egeagetete ttetgtgtte acagaatgge eatgatagge
                                                                     3660
                                                                     3720
atgetetttt cecaccact ggaaggetea caggeaaggt gagagggae acagaaggtg
                                                                     3780
ccaacactgt cgctacagta aggacctgaa gtgactttga gaaattcacc ctcacaaacc
                                                                     3840
ttccttcagg agcaggcatt ggtagtgcag aggcacagat tccgtccttt accagctgca
                                                                     3900
gaatcttggg caagttacat agcctctgtg agcctcatcg gtaaacagtg ggggttatga
aacccacctc acagggttgt tgtgaggatc caatgagttg atttaggtaa gcacctagca
                                                                     3960
4020
                                                                     4023
aaa
<210> 47
<211> 542
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (389)
<223> n equals a,t,g, or c
<400> 47
agggcacgag tttttttatg actacataat gtttattgcg atctatttta aggcttttca
                                                                       60
tggatctttt cagctatgca catggttagt cataatgatt gtcattttag gtcagagttt
                                                                      120
ctcagcctta gcattgttga cattttggtt aattctttgt tgtaggggct gtcctgtaca
                                                                      180
                                                                      240
ctgtagggtg tttagcagca tccctgatct ctacctacta aatgccagga gcaacacagt
                                                                      300
acctccagct cagttgtgac aactactgaa tgtctccaga caactactga atgtctccag
                                                                      360
acattgcaga gaaattgagt ctggttgaga agtcactgtt ttagggcata atttttgggt
                                                                      420
agactgttag attetttgtg ttegttgtnt etggeetgta taactettet taattatetg
ctactcaaat gtatttggga tcagccactg tcttccattt ctcttttgct cacagatcta
                                                                      480
ctccacagct cttctccctt caaacactgt tcctcagcat cttgtttttt gcagccaaac
                                                                      540
                                                                      542
at
<210> 48
<211> 1495
<212> DNA
<213> Homo sapiens
<400> 48
cggcacgagg ctacttatat tttatgaagg acattttttg ttagatgatc tcatcctctg
                                                                       60
tgttatttgt tgattgggtt tgttttttgc ttgttggttt gtttgtttct tccatgtaag
                                                                      120
                                                                      180
gaaaagtagt gtaagcagta ggaagaaaat gaggaagatg tattttgcat gttcttcctt
tcaatgttct tacacattgt attactgcat tgtggtaata gcttctataa aatctgccat
                                                                      240
agetgggatt atgeagettt geaagaatet actagatttt attetaacte atattagett
                                                                      300
gtcctatcaa cttctggaat tatctaatta ttgcttttaa aagtttcctg cctttcaacg
                                                                      360
tttccctgct atgcaaaacc tttcccagac cttggtttct taaaagaaag atgttgctac
                                                                      420
agttcccaat tctttcttat tacaggctca ggtgtacagg ttattctggc ttaattttat
                                                                      480
ctaatgaagc ccattccttt ttgtacatga agatgtcact taaacctatg tttacaaact
                                                                      540
aaagagacta atcactcaat atgaaaacat gaaaacattt ttgcttaaaa tattaagatg
                                                                      600
gaaatagtta aatatggatt attttgtcct tttacttttt aaaaaaagtt acatattgta
                                                                      660
tgcactgtgc tgatgcaaga attctacatt ttaatgaatt ataaaattat tctgcatctc
                                                                      720
atcacgtcac agtatttctg ctctatttat tcatatacat agaaatatat atgggcttaa
                                                                      780
```

tcatttaaaa	tttgttgcag	caagaactct	cctacctgta	ggcaatagat	tgctatgttt	840
				acgtaatagg		900
ctagaaaaat	aaagtaatgt	ttttgacatt	ggatttggtg	cagtttctaa	tgaagcaatg	960
gttggttggt	taatatgtct	tctgtagctg	ttagcattgc	caaattaaaa	agggtaaatt	1020
ttatggaaat	cctgagacca	ggaagatatc	aatttcatgt	gtacttaatg	gtataaagtg	1080
				ttatagtttt		1140
				ttcctggcca		1200
				gggggatcac		1260
				tctactaaaa		1320
				agggaggttg		1380
				atcacgccat		1440
				aaaaaaaaa		1495
<210> 49						
<211> 818						
<212> DNA						
<213> Homo	sapiens					
	-					
<400> 49						
aaaacttgag	tatgttgagg	gaaggaatat	atatatatct	gggagagaat	ggatacgttt	60
				tgtgcattct		120
				actgtgtatc		180
				ctgtccatgg		240
				cycctgaggc		300
				cgaatcccga		360
				tgcttttata		420
				gctgaagaaa		480
				tgtctgtaat		540
				cgagaacagc		600
				gggtgtggtg		660
				ttgaacctga		720
				gtgacagggt		780
		aaaaaaaaa				818
<210> 50						
<211> 1711						
<212> DNA						
<213> Homo	sapiens					
<400> 50						
ggcacgagcg	ctcctgtcct	gccactgagg	gacccggtta	ccaaccctca	tgtagctcag	60
				gactttcccc		120
				cgtttgtatt		180
				cttcactgtg		240
				tagtttccat		300
-				tcagccttaa		360
				ataaatctct		420
				agacgtgggc		480
				gacacatcag		540
				ataataggtc		600
				ggtaagctgt		660
				actgaatgaa		720
-			_	tactttaaac	_	780
agagetaa		5-5	g-ggacac	Lacoccaaac	uucyacta	, 50

```
ttcattgtat catagcttgt gatgtattct gctcatggct tttaaggtaa attgtgccat
                                                                      840
gatccactgc cattctaatt gctttaacaa gtcattacca cactactgtt acatcttaat
                                                                      900
tatgcataca gacaggtaga cttgttttac atatgtgaac taactagttg tcaaagcaaa
                                                                     960
                                                                    1020
tgcagattgt attctgcaag taaagtcttt ttctctctga aatttctagg gatgttcttt
                                                                    1080
aagtgaaatt catattaaaa ctgaagattt tagttacaag aactgagtgc agattaagtc
                                                                    1140
tttgtgattc aacatagtca agatacaact gtggatattt catggaagta tgcaataaaa
                                                                    1200
tgtctctacc tggaaaaatc tatcaagcag cgtcacagta ctgaatttga aaccagaaat
actgggtttt tatataaatg cttcatagat ttgttttatg ataaagggca cataactctc
                                                                    1260
ctaaacctca caccacctct tgaataggta taataagtcc acatcaatgc tgatgcctta
                                                                    1320
                                                                    1380
gctattatta aactcttaca gtatgatgta aagtgaaagt acaatgtaag atcattccta
                                                                    1440
qqccaacttt gaccagtttt atacagaaac atgtgccaac ttttctgttt gcaaggataa
tatcaaagca aacaccagaa agttatatct ttgatgcatt ttttcaaaat catacacata
                                                                    1500
atacacaac caaagacaaa tgatgaatat tacgtcagaa aatataaagt cttccccttt
                                                                    1560
cttcttttgc caagaaagtc caatattttc accattttta tgcacacaat caactttatt
                                                                    1620
taagctggaa gttaatgtct cattgttttc attgttctaa ataaacacct tttcccttga
                                                                    1680
                                                                    1711
gtattgctct aaaaaaaaaa aaaaaaaaaa a
<210> 51
<211> 749
<212> DNA
<213> Homo sapiens
<400> 51
gccaaaccag rtaataattt ccttataata catgaagtcg ttattttgca tttattttct
                                                                      60
taggtggcca atggggttat cttgggggga gacttttata ctcctaaggg acagcttggc
                                                                     120
cattaacttt caaagtttct ctaaagcagc gtcaggagat atatttggtt gtcatgacta
                                                                     180
                                                                     240
gtggcattcc actgacatgt aatgggtaga ggctgggtag acatcctacg atgcacaaga
                                                                     300
cagoctccca caataaagaa ctgtgtggcc caaaaatatc agtgatgctg agattgagaa
                                                                     360
acttaaagaa atttaaaaat taactctata caaaatctaa tgtttgagtt ttctccatgt
                                                                     420
atctgtgact gcaatgacca gagtgactgt ccataaagaa agtgctaaga gttggctggg
tgcggtggcc tacacctgta atcccagcac tttgggaggc caaggtgggt ggatcacctg
                                                                     480
aggtcaggag ttcgagacca gcctggccaa catggcaaaa ccccatctct actaaaaaaat
                                                                     540
                                                                     600
acaaaaatta gctgggtgtg gtggcacgca cctgtagtca cagctactca ggaggttgag
                                                                     660
gcaggagagg tgcttgaacc cgggagatgg aggttgcagt gagccgagat tatgccattg
                                                                     720
749
cggtacccaa ttcgccctat aggcagtcc
<210> 52
<211> 1091
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (1079)
<223> n equals a,t,g, or c
<400> 52
ggccagtggg cagggtcaca gggcaaggtc ccgcgggccg ctgggtgcgg cgacttccgt
                                                                      60
                                                                     120
gctcccggcg agcgggcgga gagcgggggc cgcactgggg agtgtgggct gggccgcaga
tgtcatgtgg cctgtktttt ggaccgtggt tcgtacctat gctccttatg tcacattccc
                                                                     180
tgttgccttc gtggtcgggg ctgtgggtta ccacctggaa tggttcatca ggggaaagga
                                                                     240
cccccagccc gtggaggagg aaaagagcat ctcagagcgc cgggaggatc gcaagctgga
                                                                     300
```

```
360
tgagetteta ggeaaggaee acaegeaggt ggtgageett aaggaeaage tagaatttge
                                                                        420
cccgaaagct gtgctgaaca gaaaccgccc agagaagaat taatggagga cacagggccc
                                                                        480
tatggtccta ctgtgggtgg tgacttgtcc tgctaccatg ttgacagagc cccagaaccc
                                                                        540
acatctaatt ggctttgttg cttattctgg cccttcccac accacacagc cacacaaata
ctggctgctc cttgatggcc aggcagaccc agcagcagcc gaggggccag tgaagaggaa
                                                                        600
ggccgcatct gttgtgtgt ggccacaagc actcaggcat ctgagtttac tggtgcactg
                                                                        660
ctgggaggag agttatgaga tgaacattgg ctgtcaatct ctgtgggcag geggtttggc
                                                                        720
ctctagtggg aatggctggg atttgggcgt tgcctttagg agggatacct gcatgtctag
                                                                        780
                                                                        840
ttccagtctg cactggaaag aattcaaata tgcacctggc tcccttcact attttgccct
                                                                        900
atcctttgtg ctcattctta ctgaaatctg tcttgtcagc tcaggaatgg gattccccca
ggaaggaaag cacttttctg ttctgggaag cccagactgt tcactttggg gcagggacga
                                                                        960
acatgtgcct cgtgaatttg cttgaaaaca gtcaccatct tctaccccca tcactgtata
                                                                       1020
gtgaaaaacc tgattaaagt ggtatctgag aaccawaaaa aaaaaaaaa aaaactcgng
                                                                       1080
ggggggcccg g
                                                                       1091
<210> 53
<211> 2254
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (1182)
<223> n equals a,t,g, or c
<400> 53
ggcacgaggc ccgctgcaat gttatcatca cccaacctcg ccgcatctct gctgtgtctg
                                                                        60
tggcacagcg ggtcagccac gaactgggcc cctccctgcg ccggaatgtg ggcttccagg
                                                                        120
                                                                        180
tgcggttgga aagtaagccc ccatcccgag gcggggccct gctcttctgc actgtgggta
tcctgctgcg taastgcaga gcaaccccag cctggagggc gtgagccacg tcatcgtgga
                                                                        240
tgaggtgcat gagcgggacg tgaacacaga ctttctgctg atcctgctca agggcctgca
                                                                        300
geggyteaac ceggeeetge ggetggtget catgaktgee acaggggaea atgagegett
                                                                        360
ctcccgatac tttggtggct gccccgtcat caaggtgcct ggcttcatgt acccagtcaa
                                                                        420
ggagcactac ctagaggaca tcctggccaa gttgggcaag caccagtacc tgcaccggca
                                                                        480
ccggcaccat gagtctgagg atgaatgcgc actcgatttg gaccttgtga ctgatctggt
                                                                        540
totgcacato gatgctcgcg gggaaccagg tgggatcctg tgcttcctgc ctgggtggca
                                                                        600
gagatcaaag gagtgcagca gcgcctccag gaggccctgg gcatgcacga gagcaagtac
                                                                        660
ctcatcctgc cagtgcactc caacatcccc atgatggatc agaaggccat attccagcag
                                                                        720
cctccagttg gggtgcgcaa gattgtcttg gccaccaaca ttgctgagac ttccatcaca
                                                                       780
atcaatgaca tcgtgcatgt ggtggacagt gggctgcaca aggaagaacg ctatgacctg
                                                                        840
aagaccaagg tgtcctgcct ggagacagtg tgggtatcaa gagccaatgt gatccagcgc
                                                                       900
cggggccggg cgggccgctg ccagtccggc tttgcctacc acttgttccc tcgaagccgg
                                                                       960
ctggagaaaa tggtcccttt ccaagtgcca gagatcctgc gcacacctct tgagaacctg
                                                                      1020
gtgctgcaag cgaaaatcca catgcctgag aagacggcgg tggagttcct gtccaaggct
                                                                      1080
gtggacagtc caaacatcaa ggcagtggac gaggctgtga tcttgctcca ggagatcggg
                                                                      1140
gtgctggacc agcgggagta cctgactacc ctggggcagc gnctggctca catctccacc
                                                                      1200
gacccccggt tggccaaggc cattgtgttg gctgccatct tccgttgcct gcacccacta
                                                                      1260
ctggtggtcg tttcctgcct caccegggac cccttcagca gcagcctaca gaacegggca
                                                                      1320
gaggtggaca aggtgaaagc actgttgagc catgacagcg gcagtgacca cctggccttt
                                                                      1380
gtgcgggctg tcgccggctg ggaggaggtg ctgcgttggc aggaccgcag ctcccgggag
                                                                      1440
aattacctgg aggaaaacct gctgtacgca cccagcctgc gcttcatcca cggactcatc
                                                                      1500
aagcagttct cagagaacat ttatgaggcc ttcctggtgg ggaagccctc ggactgcacc
                                                                      1560
ctggcctccg cccagtgcaa cgagtacagt gaggaggagg agctggtgaa gggcgtgctg
                                                                      1620
atggccggcc tctaccccaa cctcatccag gtgaggcagg gcaaggtcac ccggcagggg
                                                                      1680
```

aagttcaagc cc tcgaccatta ac gtcaagtcca at ctgctgctga cc ctgagcgaca gt gagctgcggc gg cttcccccca gc cgaggaccct gt ctggggctgt gt aaaaaaaaaaa aa	agggagge ggcagcgt gacgggga gacctgct gccctggg gtacagga ggcagctt acagagtg	cacacggtta cttcgtccgg cgtgcacatc gcggctggag ccgcatggtg ggagcacggg tgatgtgcgc caaatgttta	cggagccgat gactcctctc cgtgatgacg ggtgactcgc gagcggagcc cagctgcttg aagacagctg tttaaaataa	ggctgacgta aggtgcaccc ggcgccgggc gtaccgtgcg tgcgcagcga cgctactggc acgactgagc	tttcatggca gctagctgtg caccatctca gctgctgaag gctggctgca agagctgctg cctgcttctg	1740 1800 1860 1920 1980 2040 2100 2160 2220 2254
<210> 54 <211> 486 <212> DNA <213> Homo sa	piens					
<400> 54 cacactgaca tc agagaacagg aa ttttatgccc ca cgcgtggatt gt gaagcatcca tc aacttgagct tt ttttctagtc ct aaaacaggga tg ctcgag	ggaagete cateteat aatgettt atgtteet getggeaa	cctccctcct tctgtctggg tgctaccagg cgagctcagt agacagacag aaaatgactt	ctgcattttg gactccatac ggtcaggcca atgtgctccg aatgagcttg agcctcttgg	cagcctgtag gtagtggctg tactcatcac tacatgtagc aatctcagcc actttggtta	ctcacgtgcc tctaccttcc tgcaggccct acagtggaaa cagctatggc acccatctgc	60 120 180 240 300 360 420 480
<210> 55 <211> 1270 <212> DNA <213> Homo sa <400> 55	piens					
<211> 1270 <212> DNA <213> Homo sa	gataagag gataagag gataggcetgt gaggcacta gaggcacta gaggcaca gagggagcaag gagggagca gattgcac gattagtat gatagtat gatagtat gagaggag gaggag gaggag gaggag gatagtagtat gatagtat gatagag	atggtgtttc tgtctgctgg ggtcagagga gagagactct acaactgtca taaggacctt taatgggcac tttcccaagg tctattgagt caartgaaca gtgtgcaggt ctgctggcta aactcttagc actgttactg acaaattcca gccgaggtgg	ttggccatac ctgctctttc gtctgggaga gagcttcctg aaaccatgtg aggaaatagg tttatcccat tcatacagct ttaaccgaca aatgartctg gggtgtagtt taaaccttgg cttgtgcctg ataacataca tgtttggcca gcagatcacg	tgacctgtgc tgatttccca atgaggaaat ggaacaggta cctgtactat cacatcatct tttataaaca agtgatggag gatcatactg ggatttarga ggagtcagc ttgggtaagt gcacatagta aatagattgt ggcgcagtgg aggtcaggag	catttcagct caggagcagt atgagagccc taggttcttt ttggagtgct cctggataga tggaaattga gagctagcat tgttttggta cttgccagac agagttgggt aacccaaggt aatgcttaat attaatggac ctcaagcctg atcaagacca	60 120 180 240 300 360 420 480 540 600 660 720 780 840 900 960 1020 1080

gagegagaet e	cgtcycaaa	aaaaaaaaa	aaaaaactcg	agggggggcc	cgaacccaat	1260 1270
<210> 56 <211> 2059 <212> DNA <213> Homo si	apiens		·			
<400> 56						
ggcacgagcc to	cactgggta	aacacaagct	ctcggcaggg	aacaagctct	cggcagggaa	60
tgcacagaga a						120
aaaaagtttt a						180
aagcctttgg g						240
agtggtaaaa a	tacctcttt	gttatctaag	acaactttta	ggtggcatta	agaccccgag	300
ggtctagggg c						360
aggagcactg g						420
ctcttagggc co	ctgccactt	actttgctca	gagcaaaagg	tgccgtcagc	tgggtcctac	480
ccccgcctgt go						540
ctgtcaaagc g						600
agacttttct co						660
gcaggaaatt ti						720
aagatatgaa co						780
ccatgcagat as			_	-		840 900
tttgggagga ad						960
tccacatctg gg						1020
tggcgtcaag ag ttaataggaa ag						1080
caagttggct c						1140
gcagagcacc ti						1200
gcgggtgccc aa						1260
gaatcatgag ag						1320
tcgggactag ti						1380
cctaatccga ac						1440
cagtgctcag at						1500
cctgcgtctg ac						1560
tttgaggctg ca						1620
ataggaatgt go	ctcgggcag	ggaacccgga	gcaccagccg	ctgggcccct	cgttccccgg	1680
ctctgcagcg to	ctttcgggt	cctgtgggtc	tggctgtgcc	cagccctgct	gcccgtggac	1740
tatagtgtct to	gggtcctg	caggcttggc	tgtacccggc	cccagtgcat	ggtggcaagt	1800
gcctgtgatc co	cagctactc	aggaggctga	ggcaggagaa	tcacttgaac	ctgggaggca	1860
gaggttgcag tg						1920
tttgtctcaa aa					_	1980
gcccagaagt to		cttctgtcag	acacaggctg	agtatccttc	accccaaaat	2040
taaaaaaaaa aa	aaaaaaa					2059
<210> 57						
<211> 868						
<212> DNA						
<213> Homo sa	apiens					
<400> 57						
gactgactat ac	ragaaaact	ggtacgcctg	caggtaccgg	teeggaatte	caaatcaacc	60
cacgcgtccg ct						120
		J -J			555 44-	-20

```
agatggatgg tggcccaggt gcattttcca gggcctgggt tctccagatc ccgtggcttc
                                                                        180
                                                                        240
tgttgagtgg aggcaacttt gctctgtgtg aacctcgccc ctgtccctct gccgggcacc
                                                                        300
cctggcagga agcaggactc ccatcctcac cctgacttag actgtcctct gagtcagctc
ctctccaaga caggagtggg cagccttggg cagtcttctg gccccttgct aaagtgaggg
                                                                        360
                                                                        420
scaggaagct ggggctgccc tccagaaagc cggggtaggr actctgaaaa atacctcctc
                                                                        480
taaacggaag cagggytctc cagttccact tggcgccccc tcccacaagg cccttcctcc
ctgaggaccc cacccccta ccccttcccc agcagccttt ggaccctcac ctctctccgg
                                                                        540
tgtccgtggg tcctcagccc agggtgagct gcagtcaggc gggatgggac gggcaggcca
                                                                        600
                                                                        660
gaggtcagcc agctcctagc agagaagagc cagccagacc ccaaccctgt ctcttgtcca
                                                                        720
tgccctttgt gatttcagtc ttggtagact tgtatttgga gttttgtgct tcaaagtttt
                                                                        780
tqtttttgtt tgtttggttt ttgttttgag ggggtggggg gggatacaga gcagctgatc
                                                                        840
aatttgtatt tatttatttt aacattttac taaataaagc caaataaagc ctcaaaaaaa
                                                                        868
aaaaaaaaa aaaaaaaagg gcggccgc
<210> 58
<211> 986
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (592)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (669)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (767)
<223> n equals a,t,g, or c
<400> 58
gaaattaagt catttagata aaaatatgcc attcttatct gtttggtttt ttaatcttgg
                                                                         60
cttaatattt ggggttgagt catttgtttt gagagctgtc ctgtttattg cagggtgttc
                                                                        120
agcaacatcc cagatggaag cagcatcccc ctacccagct gtgacaaaaa gaaaaaaaaa
                                                                        180
                                                                        240
tgtctccaka cactgccaaa tatcttctgg gggtgcccct ggttgagaac cactgcttta
                                                                        300
gtggataaac tttaggcagg agggaaatga tcgcagttgg atagttggag gaatgtggag
                                                                        360
caagggaagc aataaactgt gaccataaaa acatagaaag atggcttata tgtggatttt
                                                                        420
tttttaaagc acgtagaatt gcttaaaatg gacaacagca gcatataaat cagtggcaga
                                                                        480
qttggtggct gaatttagag catcttaagt ctatgttctc ctggaacaga gtgcagataa
                                                                        540
ttcagttatc agcttggcta ggtgcatgtt gaagtattta gtcacacaca aacagttaat
                                                                        600
gtatggggaa gataacttct atactagtag gagagaaatg gaacaagaat wntaaataca
ctatcaaaat atgcaagaat ggcaagmgga aaaggcagaa caagctgcaa aacmcacaca
                                                                        660
caattagana taaatatttt gggacacaat aaatgtgaat ggattaaaac ctctgttaat
                                                                        720
                                                                        780
gacaaagttc tccagttaaa ggaaggcaaa tagtgttatt aggaatngat tactatatga
tgattaaagg ctcagttcaa caggaagatg attgatagaa ctttcctaca tttgtaacac
                                                                        840
agtottagaa gatattaaag caaacattca agaagaaatt gatcacctac taccatagtg
                                                                        900
tattttattg aattggtaca tttcaataaa gtgtcataag gcacggttga aggaaaaaaa
                                                                        960
                                                                        986
aaaaaaaaa aaaaaagggc ggccgc
```

```
<210> 59
<211> 695
<212> DNA
<213> Homo sapiens
<400> 59
                                                                        60
tttttttct tgaaataaaa tgggggagta atgggaaata attttttga gcccttgcgt
                                                                       120
ttctaaaaat gtttgcattg tgccttcatg tttgacagtt cagttccagg ttgaaaatta
                                                                       180
tttttctttg gaatgttaac agctgccctc tattttctgt ttttatctaa tgttgctgaa
qaqaaatctt ctgattctta ttctttttt ggtgacctgt tttaattttg tgtctttctt
                                                                       240
ttttttcccc tggaagcttt taggctctcc cttttatcct tgtagtctga gatctgacaa
                                                                       300
tgatggttgt gtctagtttt ccccaggata ctttttcatt tgtcctggtc agcactagta
                                                                       360
gatcctttca gtttgtgtac ttctgtttct cttgctctgg agaatttaaa aaatatatat
                                                                       420
                                                                       480
atttttgaga cagagtetca etetgteace caggttagag ggeagtggtg tagtetetae
                                                                       540
tcactgcaac ttctgcctcc tgtgtttaag cgattctcct gtctcagcca cctgagtagc
tgggattaca ggtgcctgcc accacgccca gctaattttt ttgtattttt agtagaggca
                                                                       600
gggtttcacc acgttgccca ggctggtgtc gaactcctga cctcagatgg tccacctgcc
                                                                       660
ttggtctccc aaaaaaaaa aaaaagggcg gccgc
                                                                       695
<210> 60
<211> 314
<212> DNA
<213> Homo sapiens
<400> 60
                                                                        60
gtcgacccac gcgtccgctt tgaggagcat tcctctagat tgcacaaggg acagtgcctt
                                                                       120
taaccaagcg aggagtccaa agctcaggac ctgactaccc tgagggcacg ctgacgcctc
ttcccagggg gatggggagc tttctgcacc cccagtggca tctcctcatc acgttctgtg
                                                                       180
                                                                       240
ccgtccttgg gaaaggcctg cattctgatc cttccaggcc cttcgagcat ggaggggcac
                                                                       300
tggggaaggt cccccgaggg aggagcacgt tgctgagtaa agaggtgtta ctcaaaaaaa
                                                                       314
aaaaaaaaa aagg
<210> 61
<211> 734
<212> DNA
<213> Homo sapiens
<400> 61
gactgcttat atttggcatt gtcttttccc tggcactgcc actgtcacca ccatcccct
                                                                        60
totggatccc tactttaccc cttcatgctg ctctggtggc agtgcctctg ctgccatgct
                                                                       120
gtacttgagc ctgctgctac agccatgcct gaagatgcag ccccttcctc tcttcctgtc
                                                                       180
ccaccaaata tgaccagete taggttecat tacttetgga etttgeteca aataaaaett
                                                                       240
                                                                       300
acacaatttt attccaaacc caggtctctt tctgcaacac ccgagaaaaa tattgggctg
                                                                       360
caggagccag agaggagag gagatttact ggtgagagtt gtaggtggga attgaaagcc
                                                                       420
aagtcatgtc tttgccccac cagaaactca ctaggatgta cacaatgcca ctgtgatggt
                                                                       480
kttaaaatat gtaactaacc tgcacgttgt gcacatgtac cctaaaactt caagtatata
taaararaga aagaactgst gatacacata tcatgaaaaa agaccaaata aaataaaaaa
                                                                       540
                                                                       600
ataaaaataa ataaataaaa taaaatatgt ccacaaatgc tttgatgttc ctttgtttct
                                                                       660
tgatctgtat gctagcaaca caggttcatt ccgtttgtga aaattcattg agctgtgctc
                                                                       720
ttatgagctg tgtacttctc tacatgtatg ttaaatgtgg acaagaactt cacataaaaa
                                                                       734
tcattttaaa aaaa
```

```
<210> 62
<211> 1410
<212> DNA
<213> Homo sapiens
<400> 62
ccgcctcctt gccgcccagc cggtccaggc ctctggcgaa catggcgctt gtcccctgcc
                                                                        60
aggtgctgcg gatggcaatc ctgctgtctt actgctctat cctgtgtaac tacaaggcca
                                                                       120
togaaatgoo otcacaccag acctacggag ggagotggaa attootgacg ttoattgato
                                                                       180
tggttatcca ggctgtcttt tttggcatct gtgtgctgac tgatctttcc agtcttctga
                                                                       240
ctcgaggaag tgggaaccag gagcaagaga ggcagctcaa gaagctcatc tctctccggg
                                                                       300
                                                                       360
actggatgtt agctgtgttg gcctttcctg ttggggtttt tgttgtagca gtgttctgga
                                                                       420
tcatttatgc ctatgacaga gagatgatat acccgaagct gctggataat tttatcccag
                                                                       480
ggtggctgaa tcacggaatg cacacgacgg ttctgccctt tatattaatc gagatgagga
                                                                       540
catcgcacca tcagtatccc agcaggagca geggacttac cgccatatgt accttctctg
                                                                       600
ttqqctatat attatgggtg tgctgggtgc atcatgtaac tggcatgtgg gtgtaccctt
                                                                       660
tcctggaaca cattggccca ggagccagaa tcatcttctt tgggtctaca accatcttaa
                                                                       720
tgaacttcct gtacctgctg ggagaagttc tgaacaacta tatctgggat acacagaaaa
                                                                       780
gtatggaaga agagaaagaa aagcctaaat tggaatgaga tccaagtcta aacgcaagag
                                                                       840
ctagattgag ccgccattga agactccttc ccctcgggca ttggcagtgg gggagaaaag
                                                                       900
gcttcaaagg aacttggtgg catcagcacc cccctcccc aatgaggaca ccttttatat
                                                                       960
ataaatatgt ataaacatag aatacagttg tttccaaaag aactcaccct cactgtgtgt
taaagaattc ttcccaaagt cattactgat aataacattt tttccttttc tagttttaaa
                                                                      1020
                                                                      1080
accagaattg gaccttggat ttttattttg gcaattgtaa ctccatctaa tcaagaaaga
                                                                      1140
ataaaagttt attgcacttc tttttgagaa mtatgttaaa gtcaaagggg catatataga
                                                                      1200
qtaaqqcttt tgtgtattta atcctaaagg tggctgtaat catgaaccta ggccaccatg
                                                                      1260
gggacctgag agggaagggg acagatgttt ctcattgcat aatgtcacag ttgcctcaaa
                                                                      1320
tgagcaccat ttgtaataat gatgtcaatt tcatgaaaag cctgagtgta ttgcatctct
                                                                      1380
tgatttaatc atgtgaaact tttcctagat gcaaatgctg actaataaag acaaagccac
                                                                      1410
cctgaaaaaa aaaaaaaaa gggcggccgc
<210> 63
<211> 1231
<212> DNA
<213> Homo sapiens
<400> 63
ggcacgagtg aatgtcgagg agttccagga tctctggcct cagttgtcct tggttattga
                                                                        60
tgggggacaa attggggatg gccagagccc cgagtgtcgc cttggctcaa ctgtggttga
                                                                       120
tttgtctgtg cccggaaagt ttggcatcat tcgtccaggc tgtgccctgg aaagtactac
                                                                       180
agccatcctc caacagaagt acggactgct cccctcacat gcgtcctacc tgtgaaactc
                                                                       240
                                                                       300
tgggaagcag gaaggcccaa gacctggtgc tggatactat gtgtctgtcc actgacgact
                                                                       360
gtcaaggcct catttgcaga ggccaccgga gctagggcac tagcctgact tttaaggcag
                                                                       420
tqtqtctttc tgagcactgt agaccaagcc cttggagctg ctggtttagc cttgcacctg
                                                                       480
gggaaaggat gtatttattt gtattttcat atatcagcca aaagctgaat ggaaaagtta
                                                                       540
agaacattcc taggtggcct tattctaata agtttcttct gtctgttttg tttttcaatt
                                                                       600
gaaaagtaat taaataacag attagaatct agtgagagcc tcctctctgg tgggtggtgg
                                                                       660
catttaaggt caaaccagcc agaagtgctg gtgctgttta aaaagtctca ggtggctgcg
                                                                       720
tgtggtggct catgcctgta atcccaacat tctgggaggc ccaggcggga gaactgcttg
agccccagga gttcagaatc agcctgggca acatagcaat actccgtctc ataaaaatta
                                                                       780
ataaataaaa agtctcaggt gaccaaaggc tcctgaagct agaaccaggt ttggataaag
                                                                       840
                                                                       900
attgaagage cacaggeeac tettecetet gageeattgg geetagtggt gteatgtatt
gtaattgctc gcagggagag cagtcttttt ggtgtaatag tgggatgtct gcttagttgg
                                                                       960
caggggttca gtccaaatgg aagaatattg ggaaataaac ctccactatc ctttatagcc
                                                                      1020
```

PCT/US99/05804

ttatttaaat	ccagtgttct	ttcataaaat ccgcagcctt ttgaactttg	ttgtctattt	atatgtgtac	caagtgttaa	1080 1140 1200
		aaaaaaaaaa		aagaaacgcc	·	1231
<210> 64 <211> 612 <212> DNA						
<213> Homo	sapiens					
<400> 64						
ggtcgaccca	cgcgtccgag	catttgtctg	tataatttta	gttattgaat	taaaatcttt	- 60
tgggacccca	acaggatgag	atcattggcc	agctggcttc	ctcccacctg	cacctggact	120
gaaattcccc	gtggcattag	aggtgtttcg	taaggtgctc	cctgctgtct	gtcctacaga	180
ttgcagtggc	tctgctggaa	aagaacggaa	ttctatgcaa	gttgcgtgtg	tcatgaaggt	240
- -		tctttgtcgt		_		300
		ggggcagagg				360
		tcttggccaa				420
		tatgcttttg				480
		tgaacccgga				540
_		caataggaat	gaaactccgt	caccaaaaaa	aaaaaaaaa	600
aagggcggcc	gc					612
<210> 65						
<211> 2270						
<211> 2270 <212> DNA						
<213> Homo	sapiens					
-220	Jup 2 - 11.0					
<400> 65						
tttttttt	aactttttaa	acaatccatt	ttaatcatct	aaattattta	caatacaata	60
acatggattc	atccttttta	agacatggga	ttgtaaaaat	caacaagtga	atgatgcttc	120
aaataataca	tttaaataca	ttaatcaaat	tttttcagtg	cttaaaactt	tttctccatg	180
ggacagcagg	ctctggacaa	aagtgcctag	catacaagtt	ttcccaattt	ccttctatca	240
taccagctgc	acataaaaag	gttcatcacc	tcctgtctcc	aaagtgtctc	cctactgagt	300
gttcccaggc	agacaatagt	tcctgggata	gtgctgtttg	gtaacagaaa	agcccaagcg	360
tagaggacgg	attaaaaggc	agggaccaga	ccgccatgga	tacaaatccc	aagacagagg	420
atgccccatg	ccttccccat	gaagcttatc	tgtctgcctg	tgtctccatg	attgcaggca	480
tagagctact	tgggacctcc	aggatgattt	acttagcgat	atgcttttta	cattctaaga	540
atcaaaatgg	tcctgtaatt	cccaatagag	aaaatagagc	${\tt caattcattg}$	ttctcccctc	600
tcccctctga	agccagtttt	taaagatgag	ccttacccag	aaaataagcc	ccaaagaact	660
ctcatctaaa	tgatcagacc	cttcctaaat	tacctttggc	aacctaggta	attcttttt	720
attacacacc	tccaacctga	ccctttctac	agtttcaact	ataaatgttc	atgcccctcr	780
	-	gaatttgcca			-	840
	_	aaggagagtg		_		900
ttctctctct	ctatcctcca	aggtaagttc	tttgttgact	tgataagctt	tagtccttct	960
gtacaacttc	tagaagatgc	acttaatggt	gcttctttgc	_		1020
gtacaacttc attctacctg	tagaagatgc taaggctgta	acttaatggt ggggagcatc	gcttctttgc ccaatcaaca	taaggcctac	ccctttagcc	1080
gtacaacttc attctacctg acgaaaatca	tagaagatgc taaggctgta gccaggcatc	acttaatggt ggggagcatc atgtttctgc	gcttctttgc ccaatcaaca accaccacct	taaggcctac gccttcctga	ccctttagcc cggacactgg	1080 1140
gtacaacttc attctacctg acgaaaatca tgctgatgac	tagaagatgc taaggctgta gccaggcatc aaaaatggga	acttaatggt ggggagcatc atgtttctgc cagtaccgca	gcttctttgc ccaatcaaca accaccacct gctggtttct	taaggcctac gccttcctga ctttttcgag	ccctttagcc cggacactgg tgtgtagata	1080 1140 1200
gtacaacttc attctacctg acgaaaatca tgctgatgac agaaataaaa	tagaagatgc taaggctgta gccaggcatc aaaaatggga aacattttca	acttaatggt ggggagcatc atgtttctgc cagtaccgca ttccctcaca	gcttctttgc ccaatcaaca accaccacct gctggtttct agcttaatct	taaggcctac gccttcctga ctttttcgag agtaatataa	ccctttagcc cggacactgg tgtgtagata ctgcctaaaa	1080 1140 1200 1260
gtacaacttc attctacctg acgaaaatca tgctgatgac agaaataaaa aaaatcaaac	tagaagatgc taaggctgta gccaggcatc aaaaatggga aacattttca cataaataaa	acttaatggt ggggagcatc atgtttctgc cagtaccgca ttccctcaca cctatgtgct	gcttctttgc ccaatcaaca accaccacct gctggtttct agcttaatct aaacaaatca	taaggcctac gccttcctga ctttttcgag agtaatataa catgacttga	ccctttagcc cggacactgg tgtgtagata ctgcctaaaa tgacttctct	1080 1140 1200 1260 1320
gtacaacttc attctacctg acgaaaatca tgctgatgac agaaataaaa aaaatcaaac aaaattaatg	tagaagatgc taaggctgta gccaggcatc aaaaatggga aacattttca cataaataaa tcaaggaaaa	acttaatggt ggggagcatc atgtttctgc cagtaccgca ttccctcaca cctatgtgct aaggaaaagt	gcttctttgc ccaatcaaca accaccacct gctggtttct agcttaatct aaacaaatca tgatcccaag	taaggcctac gccttcctga ctttttcgag agtaatataa catgacttga taaaatccct	ccctttagcc cggacactgg tgtgtagata ctgcctaaaa tgacttctct tgaccacagc	1080 1140 1200 1260 1320 1380
gtacaacttc attctacctg acgaaaatca tgctgatgac agaaataaaa aaaatcaaac aaaattaatg	tagaagatgc taaggctgta gccaggcatc aaaaatggga aacattttca cataaataaa tcaaggaaaa	acttaatggt ggggagcatc atgtttctgc cagtaccgca ttccctcaca cctatgtgct	gcttctttgc ccaatcaaca accaccacct gctggtttct agcttaatct aaacaaatca tgatcccaag	taaggcctac gccttcctga ctttttcgag agtaatataa catgacttga taaaatccct	ccctttagcc cggacactgg tgtgtagata ctgcctaaaa tgacttctct tgaccacagc	1080 1140 1200 1260 1320

tgtatttcaa	ggggagtgtt	gtggccttca	caaatgaaaa	ttatgaatca	caaagataaa	1500
			ggaatgtcat			1560
aattcatttt	ttctcagtct	ccccctgac	tccacaaaag	ctttgccttc	ccaacacaag	1620
gggctgggag	gtccagtcta	gacagagcat	gctgttgggg	taaacagtaa	ccatgtgatc	1680
ccatgattcc	cagagetetg	agcacaaagc	ttttcatccc	agtggcaact	ggaatgtggg	1740
			atgcttgggg			1800
			ctctctggaa			1860
			aacacttttt			1920
			acctccattg		_	1980
			cctggttcta			2040
			cccagcactc			2100
			gaatcttccc			2160
-			agggacccaa			2220
-			tcgataccgt			2270
		_	-			
<210> 66						•
<211> 1283						
<212> DNA						
<213> Homo	sapiens					
<400> 66						
ggcacgagcg	agggaacaga	ttcccagggt	ggggtggggt	agggctgggc	ctttgctcct	60
ttgtctcctt	tcccaaggca	aagtgaagag	aagagagtga	ctccttcttc	actcagggaa	120
cccaggcagg	gatgaagcgc	ctgtggtgtc	tgagctgggt	cccggggctg	caggggagcc	180
cctcagtgtt	gtcctctgta	ttcttctccg	tgttcaaacc	acagctgcat	tggacatgca	240
gtcaggtgtc	ctctcactgg	caccctccct	gccttttcat	tctttttct	ggatagtctt	300
tcatcaagtt	ttctctgcct	tcaccttgct	cttcctgaat	cagttcacct	tgaggggggt	360
taacagagca	ccttggcagg	ctctgttcct	ccaggtccca	ggccagcccc	cgggactcag	420
ggcctgcctt	cccctcacct	tcttgagcag	cacaaactcg	ttctctgctg	ctgtccgctt	480
gttgatttcc	tcctcatacc	tgtgggaatg	gcgagggctc	attggttaat	atctcactga	540
aagcccttgc	tttcacaggg	cagcgttgag	caaggagcag	cgtgtccatc	agaagataca	600
ggtgctgggg	gccgcagagt	actgggcagg	ggtaagtggg	ggaaggcttc	ctggaggagg	660
gaacatgcta	accagtttgg	aagaatgaga	ctgttaaaga	tccagcttgg	caaacgagga	720
aggagcacat	ggagcgaatg	ccctgtggga	ctctcagagg	aaccaggatg	tgaactgccc	780
tccccaaatt	tgagtacagc	tttacaaatg	acaaagcgct	ccaatctgca	tttcctcagt	840
tacccttgag	agcagtcttg	gagccagatg	cacttaaccc	tcctcttaca	tgggagagaa	900
cgtgagtggt	ttccccaaac	attccctaaa	cccagagcca	gagataaccc	tctgcccact	960
gcccagctca	ctgggcattt	gtcctaagag	tcaġgccaga	ggctggagga	gcagagagca	1020
agttccagag	ttttgttggg	gtgactctgc	ttgatatgac	caagaacaat	gccctccact	1080
			gccaggggtc			1140
		_	cacctcttct		-	1200
cgcacattcc	tcagctccga	gtccagcctc	accctgtccc	cagacagcgt	ctccagctgc	1260
ttccgcaggt	tgctgatgta	gcc				1283
<210> 67						
<211> 1263						
<212> DNA						
<213> Homo	sapiens					

<220>

<221> SITE

<222> (1256)

<223> n equals a,t,g, or c

```
<400> 67
gaggagateg ccacctccat egaacccate egegacttee tggccategt tttettegee
                                                                         60
tccatagggc tccacgtgtt ccccacgttt gtggcgtacg agctcacggt gctggtgttc
                                                                        120
ctcaccttgt cagtggtggt gatgaagttt ctcctggcgg cgctggtcct gtctctcatt
                                                                        180
ctgccgagga gcagccagta catcaagtgg atcgtctctg cggggcttgc ccaggtcagc
                                                                        240
gagttttcct ttgtcctggg gagccgggcg cgaagagcgg gcgtcatctc tcgggaggtg
                                                                        300
tacctcctta tactgagtgt gaccacgctc agectcttgc tcgccccggt gctgtggaga
                                                                        360
gctgcaatca cgaggtgtgt gcccagaccg gagagacggt ccagcctctg atggctcgga
                                                                        420
gatgatggac cgtggaaggg aagcgtctgt ggggagtgag cgcttagatg gccagcagct
                                                                        480
                                                                        540
gctccttctg ggaagctcgc accttggcaa cagaacagcc ctctagcaga gcgtcagtgc
                                                                        600
agtogtgtta tocoggottt tacagaatat tottgtocta tittagaatt ticoggagta
gtttatttgc agtctgttga ttatgtgcag tagacccggg acactgcgtt ttaccgatca
                                                                        660
ccttgaatgt ggtgcctgga tgtgcctttt tttttttcc ctgaaattat tattaatttt
                                                                        720
ctattgtgag ttcatcagtt catagttttt ttagtaaaga agcaaaatta aaaggctttt
                                                                        780
aaaaatgtac aacttcagaa ttataatctg ttagtcaaat atttgttatt aaacatttct
                                                                        840
gtaatatgaa gttgtaatcc tggccgtgag cttggaagct tacttttgat tcttaaagcc
                                                                        900
tatgttttct aaaatgagac aaatacggat gtctatttgc cttttattgt aacttttaaa
                                                                        960
tgaaataatt tcatgtcaat ttctattaga tatatcactt aaaatatttg gttttaaatc
                                                                       1020
acaagaatat gtattettta ataaagataa tttatgatea tggtataatt aattgaaatt
                                                                       1080
tattaaaatc tgtttttatt aaaaaaaaaa aaaaaaaaac tcgaggggg gcccggtacc
                                                                       1140
caattcgccc tatagtgagt cgtattacaa ttcactggcc gtcgttttac aacgtcgtga
                                                                       1200
ctgggaaaac cctggcgtta cccaacttaa tcgcctttgc agcacatccc ctttgncagc
                                                                       1260
                                                                       1263
tta
<210> 68
<211> 1617
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (1578)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (1586)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (1605)
<223> n equals a,t,g, or c
<400> 68
tegacecacg egteegggaa acetgatact geaacetgea atgtaggatg tttgtatgge
                                                                        60
atttaaaggt aatggtgatg tttattattc tatactttgc atactgtgag agtaattttc
                                                                        120
actetytett aagtgtgagt aageetette taaaaatett gttettgeea agaaatttat
                                                                        180
aaatcacata cgaagacgtc tgttgctaac agttaacttt atgaggtaac tatatccttc
                                                                       240
tatttctctg gactcatttt taaaaaaatat gccgaatact gcatactgtt taaggtagta
                                                                       300
tataagttta tgagagaagt ggagagcttt cttccttgaa aagtcggtat ttgttgagat
                                                                       360
accatttgcc tcacagagag gtgttcccca ctcccatccc cattgccaga taaataaata
                                                                       420
ttttgagaaa agtgacctaa aacagctgga aatcttaggt gcatctgtct gcagacctcc
                                                                       480
```

540

39

ttaagcaggc tgtatcttac aattccctta ctgcactggg taagtgttaa cttagttttt

```
qttttgccct ttgctttaaa tattctccaa attaccattt atgcaacatg gttagggtta
                                                                    600
atactgcatg gtattcattt acttgtttca tgaactttcc agtactgtac aaggtcaaca
                                                                    660
aagtaatgcc tgtggtatcc tcatctctca cttttttact ctgtggtttt agcacagtaa
                                                                    720
                                                                    780
qqtactgcaa agaccttcct tccaaatgtc tccttgactt tattccttgg gccaattcag
tatcctcaac atcctaagat tttgttgttt tatcactgac ctgtggttgg cctgttttat
                                                                    840
tctaatttcc agaaaagttc aatcccagta tttgcaatat caaataactc taaaaccgat
                                                                    900
gttgtgattc taccttcctt actatttta ctgggcaaat gccctatttt tttaattatt
                                                                    960
                                                                   1020
attatttta acttttggga cacacaaaaa tcagcaattc tcatgaagcg tttgttagtg
tggcagactt gtctaattcc tgaaactcat tcatccctt gagccagcca atggggagga
                                                                   1080
ataggataat gcaaacacat gttttgtttt ctcattttca aataatttac catgttaaaa
                                                                   1140
                                                                   1200
taaacttttc tttgtttttt atttgtagag tcagctaagt acccatattt aaatgccgtc
                                                                   1260
aaggtaactg ggcaatcaaa caccttttgg ggattctggc tttagtattt tatcagccat
                                                                   1320
                                                                   1380
tttaaaatta aatataaaaa tootttgtaa gaaacttgca tootaatttt totttattgo
                                                                   1440
aattgaaagt gtaaataata agacaatgta agtaagacct tcctaatgtc taatacaaac
                                                                   1500
tgccgtgtgt gtggttttct tttttaaatg gatagtagag tggtggctgg ataagggtac
                                                                   1560
                                                                   1617
ctgtaatggg ggtttggnca gcaagnctga aatttatact tttgnaaata aaactac
<210> 69
<211> 1389
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (755)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (1177)
<223> n equals a,t,g, or c
<400> 69
                                                                     60
gcttttttag gcattcattg gacacttgct ttaataagaa tagtttctta gttggcataa
                                                                    120
tgcttctgaa gtaatggtac tttaaaataa tttatctcat gaactttaag cattcctctt
gaaatcttgt ttttactcta tctagtcagc ccttataggc aatccaaagg gatgctttgg
                                                                    180
                                                                    240
atgctttagt ccagtggttc tcagagagtg gtctgtggag tcctggaagt cgctgagacc
ctttcaggca atttgcaagc tcaaaactaa tttcagaatg acaccaggat gttctgtgcc
                                                                    300
                                                                    360
ttttttgctg tgttggctat ttgcactgat gatgcaagag aaatggggag gagtaaaatc
                                                                    420
actggtgtct taccactatt caagacagtg gcaccaaact gtagtagtct agtaggcatt
                                                                    480
gtatcctgca cagcctcccg ccaggagtgt gatggaagac caaggaagca gtgtaaatgg
                                                                    540
aggtacacgg gaagcacttc ctttgcaccg tgaaggatga tgactgtttt aaggaaaagc
                                                                    600
acttatgcca tggtttgtgt cgtgagctga accagctgcc tttttcatgg aacatgactt
                                                                    660
ttacttgaaa gagtaactga cagaaaacca attattctga cttgtgtttg tagcagacat
tttcttgaaa atgaagaagt gggtctgtga cttcaggaaa atgtctgaca gtatctgttg
                                                                    720
ccaaagaaaa tgtgagtttt caagccaaga atagnaatct agaaaacttg tattcmcctc
                                                                    780
catgggcttg atgmcctctt agtacttaga cctttctgac gagataagcg gtggcattaa
                                                                    840
caaatgtgac gtttttcatg ttatctaaat acatttgtca acatttrraa gatctgcaca
                                                                    900
                                                                   960
actccctgga ccaggattty ccaaatgatt gttgcttttt gttacaaaat cagggaatag
atagaagatt cattcaaata atgaaagata gactgatgga tttttatgta acagaatagg
                                                                   1020
aaaagtttat gacatgtttt cagattccac cttgcaacta atttttaaga agctaccact
                                                                   1080
```

```
tgtcagccag gcacagtgct cacgcctata atcccaacac tttgggagtc caaggtgggc
                                                                     1140
agatcacttg aggtcaggag ttcaagcctg gccaacntgg tgaaagactt ctctactaaa
                                                                     1200
actacaaaaa ttagctgggc atggtggtgg gtacctgtaa tcccagctac tctggaggct
                                                                     1260
gaggcaggag aattgcctga gcctgggagg cagaggttgc agtgagccaa gatcgcgcca
                                                                     1320
1380
aaactcgag
                                                                     1389
<210> 70
<211> 1896
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (1802)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (1856)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (1886)
<223> n equals a,t,g, or c
<400> 70
aaaacaaaaa agctaataat ctcctcaagc aatttctggc ctaataqaat tataqtaqac
                                                                       60
agtgaagtat ctaaacccag ggaatcagat tgaggcacca tgtccatcgc cttgagaatt
                                                                      120
aataggctgc atttctgggt tctccttttt ttttttttt ttgcccaact gagtctttct
                                                                      180
gtggacttac atggaacttc ttattctctt aaatcattaa gttacttgac aatattcttg
                                                                      240
gatttggaga aactggatgt agggccgtat gaaaaaatca ttcgaaatca gatttagggg
                                                                      300
tataaggttg gataggaatg ttttagaaag aagaatgtaa ggcagataac taatttgtca
                                                                      360
catecaaagt ataaaactgc tactttttcc ctagaaaagg gaagctcatt ttaggcagcc
                                                                      420
taaaccagta agattttctt cctcctccaa gtgcagattt ttgtaccttt cgtttgtcaa
                                                                      480
aacattcttt ggccctatgc atgccagagt gatatagaaa ggaagttacc acatttttt
                                                                      540
gagaacaaat cactcctgat aaaatttctt agacaattga taatcatttt aagaagaaat
                                                                      600
ttaattgtat ttagctctgt gtctcgcccc tttggtgtca ctcttctacc tcttccatca
                                                                      660
ctatagctaa atatttagaa gtatatcttg acacctagca caaatgtttt ggttaagtat
                                                                      720
cttaaaactg atggatggta tggctggggc agcatggctc acgcctgtaa tcccagcact
                                                                      780
ttgggaggcc aaggegggtg aatcacctga ggtcaggagt ttgagaccgg cctgaccaac
                                                                      840
ttggagaaac cccgtctyta ctaaaaatac aaaaattagt csggtggtgg cgcatgcctg
                                                                      900
taatcctgtc tactcaggar gctgargcag gagaattgcc tgaacccggg argcarargt
                                                                      960
tgcartgagc tgaratcgtg ccattgcact ccagcctggg caacaagagc aaaactcagt
                                                                     1020
ctcaaaaaac aaaacaaaaa acctgttggt atagtacgaa agaaacgtct tgcagttttc
                                                                     1080
tgttgcagag aattaattag aaccaacctg ttggattata cacattcacc tttcagaatc
                                                                     1140
ctttcttctc tgtggaaacc cacactctca gcagtgtgtg ggaacacagt agattcttaa
                                                                    1200
ggaatgettg ttgaatgttg cagtetgeat ettettgaag taacagaact gttggtaget
                                                                    1260
gtttaaaagt aaaatgtgtc taaagacctt ttggaaatta agatgtaaga gattaatgca
                                                                    1320
ccaaagcagt ctcttaatta cttaaaatga attatttcaa agaatcttta attgaatttt
                                                                    1380
ctgtgaagtc tggaatttgt aaattatgtc cctttgttca aaccagcccc tgaaaagaac
                                                                    1440
aattaaggca attaagatag cattaaagtt ttcaatgaag ttggcatttt cygtgtatta
                                                                    1500
agattagatg ttagctgctg aagtttgtgg aggtcggaca taaagcttcc aacatcagta
                                                                    1560
```

aaagggccat tttgatattc attttgaaaa angaggatta	gtagcatgcc atycaagttc gttttccaaa aaccttccag	tgcgataaaa tcaaagccag aattttcyca gaaagtaaaa gttccaaatg gtaatnaaat	gttactcagc cctgatttwa aatttaaata gttttgggtg	ctagtccttg kgattaattt atccggtaac	tttaagcagt cctkggaaaa cccgtataat	1620 1680 1740 1800 1860
<210> 71 <211> 308 <212> DNA <213> Homo	sapiens					
cgggcgccgt agctggagct ctgatgaggc	gctggccgcc ggacgtggag gggtcggccg	ggacccatgc agcctgctct ctggtgcccg ccacccgagt gggacattaa	gggcgtgcgc aggacgacgg gagcgacacg	cgtgggcctc gacggcctcc gccgtggggc	tacatggggc gcggaaggcc ctggcaggcg	60 120 180 240 300 308
<210> 72 <211> 1688 <212> DNA <213> Homo	sapiens					
<220> <221> SITE <222> (912) <223> n equ) uals a,t,g,	or c				
<400> 72						
		ggacttatgc				60
-		tttcctctcc				120
		aaggagagaa				180 240
		ggacaaagat taacgattaa				300
		actctgcaag				360
		cctactagtt				420
•		aaatgtacct	=			480
		tctcttgcct				540
		aaaaagagaa				600
aatagatttt	aaatggctaa	actactagec	ttaaaactac	tagtctataa	atcaactacc	660
acttttgtga	atctgacagg	ccacatttt	atatggccct	ttacagaatg	gagtgtgttg	720
	_	gagttgagct				780
		gcaacagatc				840
		aaacaagctt				900
		gtttgggctt				960
		gatgatattc				1020 1080
		tgaaaactgt tgtgggtgag				1140
		ctccaatttt				1200
		ttaccttcct				1260
		atttagttca				1320

```
1380
gcctgtggta tcgaggtgtt gcaaacaata tctgtgttgc gcttcctgtt ttaacctacc
tegttttgtt tgtttttgtt teaetgttea teaeageagt gttateteea ggagacatat
                                                                     1500
agagagctca accggcaatc tcaggtgcat ttaacatttt taaaacgaaa cagtagttga
                                                                     1560
ccaaattttt cttcttaaaa aattggaagt ggggggaatc caatgacaaa aactaatgtg
                                                                     1620
gcttgtttct ggagaaaata attactgtaa atggaacaac aacaacaata aaacacacgt
                                                                     1680
taaacatc
                                                                     1688
<210> 73
<211> 1138
<212> DNA
<213> Homo sapiens
<400> 73
gggcgcctgt agtcccagct attcaggagg ccgaggcagg agaattgcct gaactcagga
                                                                       60
ggcggasttg cagtgagccg agatcgcgcc attgcactcc agcctgggtg acagagtgag
                                                                      120
actetttete ecaaaaaaaa aaaaaaaaaa aaagteaaat geagetggga atgtggtteg
                                                                      180
tgcctttttg tatattaacc atttgaaact tggttgtaag gtggggttgg caatgtcagg
                                                                      240
cctggctgca gcagctcatg tctttagagt gtgcctcttc cctctctcgt ggggctcgag
                                                                      300
caagactacc ttcatacatg ggctctccag ttacatagca actccagtgt taaattccat
                                                                      360
cttttcttcc tggaaaagcc gtagaaagga cacctggaca tgcctgctgc acaggttgtc
                                                                      420
tgccttcccc atcagccsca gaaggaggaa ctttgctctc ttctctcaca gctgtgtgt
                                                                      480
cataagaagt agttcggatg atgtgggtcc caccatgtat tccttctctg ttccatgtag
                                                                      540
agtaaaataa atgggagttc tgtttaatgc atcacctcgg ttcatattgc atttqccaag
                                                                      600
aaagtgcaat titattgaac attaggattg aattettaac tgagtaatea attteagtag
                                                                      660
taagttaaaa tgccttctat taatggacaa ctgcaaccgt taatcagagt tacagtagat
                                                                      720
taacagttgt cagcatttat gctaatagca ctaataaacc gtgggctcat gatttgcact
                                                                      780
ttataattcc atatttctca aaacagttgg taatactttt tgcttgaagg tattgattct
                                                                      840
tttgtccctt tgcttgctac ttggagatgt agagaaagct aaatgacatt ttcacggtga
                                                                      900
tgacacaata tcaccttctg cttttgcaca cttggctttg tgtcaaaata gatggaaagg
                                                                     960
gttcatttgt tctggtgctc tactgtttaa tttgatctgg tgtgtgacta aagcaagaca
                                                                     1020
aatagtattt ttaatgaaac catttaataa cctctggtag cttagagtcg aaggcattgg
                                                                     1080
aaaaatgcaa ttaaaggatg cctagatgta aacaaaaaaa aaaaaaaagg gcggccgc
                                                                     1138
<210> 74
<211> 777
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (761)
<223> n equals a,t,g, or c
<400> 74
gtagcacctt gaaattgggc agttctgaat atgctgttga gtaaagggaa aatcactatc
                                                                      60
tttttaggac ctttggaatg tggctccatg catttgctaa cgttgttctc ttcagggctg
                                                                     120
atttttctgg gctgttctac tcctctatcc ttctgtgatt gtcttccaat tcttttatta
                                                                     180
tggttagagt tccctgtaga aaccagtggg gtgtgtagtt aacaagtgtc aaaaggagta
                                                                     240
gaataattac tttatgtgat gtacttacga gaatactact tagtagaatc cagtataacc
                                                                     300
aaacaaaatg tagacgtatt taactatcca gtgtgtcagc tacacatttt ctcatcttta
                                                                     360
tcacttctgc tctggatatt gtgatctact ttcctatctc tttgtatttg tkttttcaca
                                                                     420
tttctttttt gtgtaaactg ctctatactc ttattgaaac ttgagcataa ttttatattt
                                                                     480
```

```
acatagtaaa gtttctgata tgattagaac ataagttgtw tctcctaatt ttccaataga
                                                                        540
gccattgatk tttcatttta acccctttta atggaacact tactagcttt ctaaattata
                                                                        600
aacttactaa ctttaaatat acatgaatta tttctcagca ttcttaaaac aagggtactg
                                                                        660
aacttcgwtt tcttgatgat tcaggggaaa agaattctga gtgttggaaa ggactttaga
                                                                        720
tgtcttccaa ggatttatgg gatcaattta agaaaaaaa naaaaaaggg cggccgc
                                                                        777
<210> 75
<211> 1060
<212> DNA
<213> Homo sapiens
<400> 75
gatgtatttc cttaatatgt agtttcagaa gtggaattta ttagagttaa actaaactca
                                                                        60
ttaaatttag agtttcttat tgtctttcat gagaacattt ttccttttca ttcataaatg
                                                                        120
atattgaaac actatattct tactttcata ttcctgttta tatttttgtt tttcatgtta
                                                                        180
aacattttac attctaatag taacctcatc gacctgttaa aaggcaatat aagatttaga
                                                                       240
ttattaaata gcatgtaata tatgtgatca gtaatcttca atgagcttgk tcttcattta
                                                                       300
attgcaacgt tatgtctgat tttttttgkt gcaaagcttt cagaatcttg acttgtggta
                                                                       360
atcttctttt aaaaaagctt ttaacagaat taatargtca tcacgttatg ataaatgatt
                                                                        420
aaggaaatga tgcctctaat acatkgaatt attaaaacta tcattttgaa aaattatatt
                                                                        480
ggtacaaact agtgtctact gctattactc atacatttca gaattcatac atggatatcg
                                                                       540
tctaggattt ttttttgcg taatcatgag ttacggtgtt aaagttatag tgttaattta
                                                                       600
attatgttat agtgttaatt tatctgtttt acatctcact tttgtatctg aaaccgttcg
                                                                       660
aaaataatta ttattaaagg ccagttgaca aaatttccac tcctcctccc cagtgtgact
                                                                       720
ttccttattt gtattatacc tataaagact acctcttaca tcggccaggc acagtggctc
                                                                       780
acgtctgtca tcccagcact ttgggaggac gaggtgggca gattgcctga gctcaggagt
                                                                       840
tggagaccag cctgggtaat atagtgagat cctgtctcta caaaatatac aaaattagct
                                                                       900
aggegtgeet gtagteecag ctaettggga ggetgaggtg gtaggatgge atagagteea
                                                                       960
ggaggcagag gttgcagtga gctgagatgg tgccactgca ctccagcctc agtgtcagag
                                                                      1020
ccagtccctg tctcaaaaaa aaaaaaaaa gggcggccgc
                                                                      1060
<210> 76
<211> 1503
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (6)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (18)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (41)
<223> n equals a,t,g, or c
<220>
<221> SITE
```

44

<222> (1501) <223> n equals a,t,g, or c gtggangccg ctcctganaa ctagtgggtc ccccgggctg ncaggattcg gcacgagaat 60 gaatggcaaa gaaatagaag gggaagaaat tgaaatagtc ttagccaagc caccagacaa 120 gaaaaggaaa gagcgccaag ctgctagaca ggcctccaga agcactgcgt atgaagatta 180 ttactaccac cctcctcctc gcatgccacc tccaattaga ggtcggggtc gtggtgggg 240 gagaggtgga tatggctacc ctccagatta ctacggctat gaagattact atgatgatta 300 360 ctatggttat gattatcacg actatcgtgg aggctatgaa gatccctact acggctatga tgatggctat gcagtaagag gaagaggagg aggaagggga gggcgaggtg ctccaccacc 420 accaaggggg aggggagcac cacctccaag aggtagagct ggctattcac agaggggggc 480 540 acctttggga ccaccaagag gctctagggg tggcagaggg ggtcctgctc aacagcagag aggccgtggt tcccgtggat ctcgggggcaa tcgtgggggc aatgtaggag gcaagagaaa 600 660 ggcagatggg tacaaccagc ctgattccaa gcgtcgtcag ccaacaacca acagaactgg ggttcccaac ccatcgctca gcagccgctt cagcaaggtg gtgactattc tggtaactat 720 ggttacaata atgacaacca ggaattttat caggatactt atgggcaaca gtggaagtag 780 acaagtaagg gcttgaaaat gatactggca agatacgatt ggctctagat ctacattctt 840 caaaaaaaa aattggctta actgtttcat ctttaagtag cattttgctg ccatttgtat 900 tgggctgaag aaatcactat tgtgtatata ctcaagtctt tttatttttc ctcttttcat 960 aaatgctctt ggacattatt gggcttgcag agttccctta ttctggggat tacaatgctt 1020 ttatcgtttc aggcttcatt ttagcttcaa aacaagctgg gcacactgtt aaatcatgat 1080 tttgcagaac ctttggtttt ggacagtttc atttttttgg atttgggata gattacatag 1140 gagtatggag tatgctgtaa ataaaaatac aagctagtgc tttgtcttag tagttttaag 1200 aaattaaagc aaacaaattt aagttttctt gtattgaaaa taacctatga ttgtatgttt 1260 tgcattccta gaagtaggtt aactgtgttt ttaaattgtt ataacttcac acctttttga 1320 aatctgccct acaaaatttg tttggcttaa acgtcaaaag ccgtgacaat ttgttctttg 1380 atgtgattgt atttccaatt tcttgttcat gtaagatttc aataaaacta aaaaatctat 1440 1500 1503 naa <210> 77 <211> 872 <212> DNA <213> Homo sapiens <220> <221> SITE <222> (844) <223> n equals a,t,g, or c <220> <221> SITE <222> (858) <223> n equals a,t,g, or c <400> 77 ggggaagttc ttcactgcct tgcatttgac tccagatccc tccatcctcc cagagccttg 60 gcctcaaaaa tgctgattct agcatcatgg aaatgctgtc ctcaaagtgg tctaaacggg 120 ttgctgcttc acttgctcac ttaatctccc ttttcatagg gctgttgttt ttacttctgg 180 gaagttctgt ttaccctgga acagaaactc tcttccctaa aagttgattt tattgaccca 240 tggaggccag agacacttag gcatattttc cctccagact agaagcttct gaggaggacc 300 teetgagtet geaccetgge teeetgetgt getgagggee eeegtgttaa ceteaegttg 360 tgcctcctct gattcagagg gcccagtgtg gttctgtcag ccaggcagtg gccccagctc 420

```
tacagaaatg agttgtcatt gcatcctagg gccagggtct tcgtgcttgt gtgtgttacg
                                                                     480
tggaagtatg tggacaccaa gtgttcctgg atggccacag cctgcgaagg aaactggggc
                                                                     540
cagcagctgc tctgtgtttt cagccaacaa tggctcctgc ccactgccgc tgcataacca
                                                                     600
ccagaggcag gcttctcttg acacaggcct gtcgttggag catgtgcctg gcgagtccta
                                                                     660
tttctattcc cctgtgggtt agggacaggc agctgtacct tcagtgtgtt gctggggcag
                                                                     720
780
cagtctgcag gacagagaga ggctmtatct caaaaaaaaa aaaaaaaaa actcgagggg
                                                                     840
gggnccggga cccaattngc catataggaa aa
                                                                     872
<210> 78
<211> 573
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (560)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (563)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (566)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (567)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (571)
<223> n equals a,t,g, or c
<400> 78
gatcaagttc cttagttttg acaatcaggt cccaaactct ttttcttgcc tcatttattc
                                                                     60
attcaacaag tattttttgt gcactgaata tgtggccctc actaggcaga tgctgcctat
                                                                    120
tcttttgcct gttaactaat ttaacctctt gtcatacctc ccaaatcacc ttatgctcca
                                                                    180
gagaaacttg tgtatggtca cgtaccacat aatgatgctt tggtcaacta cagatagttg
                                                                    240
atccagtaag attaccatgg agctgaaaaa ttccaatggc ctagtattta ctgtgctttt
                                                                    300
aattattatt ttggaatgta ctccttttac ttataagaaa cttagctgta aaacagcctc
                                                                    360
agtcattttc ttcatgaggt atttcagaaa atggcattgc tatcatagga gatgacagat
                                                                    420
tcatgcattt tattgcccct gaagactttc cagtgggaca agatgtggag gtggaagaca
                                                                    480
gtgatattga tgatcctgac cctgcataaa tctagcttaa tatgtgtgtt tgtcttagtt
                                                                    540
gctgacaaaa aaaaaaaaan aanaannaaa naa
                                                                    573
```

<210> 79 <211> 1509

<212> DNA						
<213> Homo	saniens					
\215> 110111C	Jupiens		•			
<400> 79						
	tgtacctaat	gagettetee	attcacttto	taaaaataat	ttgtatgtgt	60
					tcccaggcca	120
					ccaagacctc	180
					gtgaaaacaa	240
					tgtccccaga	300
					gcccatcagt	360
					ctcaggggcc	420
					gacttttaaa	480
ggaattatto	tggcagcaca	tgtagtattc	ttggatgatc	ttgctgctct	tatttctcct	540
					catctgctta	600
					ggatgtagag	660
agaaatagga	cttaattcca	ctaggggctc	tcatctcaca	ccttaaggag	gagatttcta	720
gaaaaactgg	gccagatttt	ctttgttctc	catcatttta	atgtggcagg	ctgttcagtt	780
ttcttactct	tacctatgtg	atatttcttc	gtaacgtgtc	caaaaagaaa	aaagacccaa	840
tcagtgtctc	ttgactttgt	tctttgatcc	ctcagtttct	tcttgatttc	agcatgtgtc	900
gggttcctaa	ttttgggtat	gagttagcaa	atttaaccat	tgtgtttgtg	ccctacccag	960
gggactcccc	agtttctgac	ttgaagtaga	ctgagaagaa	tccacgaggt	gctatctggc	1020
cagatttaag	tagattctat	ttccttggtt	ctccctctcc	ctgaggacct	cttattttat	1080
tgtcccctct	tctaggttaa	ttctcctttg	atttgacttt	gttgagaagg	aggttggaca	1140
gtagattagc	aaagttccaa	gtgcaaaatt	acagtgtgtt	agagtgtggg	gggaaaatta	1200
gtcttatttt	tccctacatg	ggatacaaca	ctgtgaattc	aatcttcaac	tgaaggccct	1260
gcagttctcc	taaaacatag	ttgtttgttt	ttctttaaca	aagtttaagc	tagtgttaat	1320
	aaattgcttg					1380
	tgtaattcat					1440
	ttataaagag	ttcaattgta	aataaactat	tgtggctgtt	aaaaaaaaa	1500
aaaaaaaa						1509
<210> 80						
<211> 1109						
<212> DNA						
<213> Homo	sapiens					
	_					
<400> 80						
ccacgcgtcc	ggccgcagaa	cgggctccgc	ggacgacggg	ctccagggac	gcacaggcag	60
	accgcgggtg					120
gtcgcggcct	cggccgctgc	tgctgctgtt	gctgctgctg	tcgccttggc	ctgtctgggc	180
ccaagtgtcg	gccagggcct	cgccctcggg	gtccctgggc	gccccggact	gccccgaggt	240
gtgcacgtgc	gtgccgggag	gcctgccagc	tgtcggcact	ctcgctgccc	gccgtgcccc	300
	cctgcgcctg					360
	cttcgcggga					420
	gcatgtgcga					480
	gctggaagca					540
	ggccggcaac					600
	ctcactcagc					660
tgggccgcct	gcccgctcta	gacgcgctgc	acctgcgcgg	caacccttgg	ggctgcgggt	720
	cccgctctgc					780
	ctgcgtgtgg					840
	tagccattgc					900
	cggcctcctt					960
cicacogodt	gccgtgcgcg	cegeegeege	cccccccc	ecgecetecg	cccgccgaga	1020

	ccgaaccccg cgctgccgcc		ccacggctgt	gcctcgcccg	cggacccggg	1080 1109
<210> 81 <211> 807 <212> DNA <213> Homo	sapiens					
<400> 81						
	cggacgtcct	gatagatect	ctgctccaat	aggcaactcc	ggccttccct	60
	ggaacctctg					120
	cctgtggagt					180
tgcttgcggg	tgccgggagg	cctggcctcc	cccagggccg	ccacctctgc	tggttgctct	240
gtgctttcac	cttaaagctc	tgccaagcag	aggctcccgt	gcaggaagag	aagctgtcag	300
	aaatttgcca					360
	taatttccgg					420
	atgcagctca					480
	cgcttatttt					540
	tgtcatcatt					. 600 660
	cgagtccata agtgaaagtg					720
	ccttttcccc					780
	aaaaaaaaa		cccagaccac	caacgagcag	uucuuuuugu	807
J						
<210> 82 <211> 1043 <212> DNA <213> Homo	sapiens					
<400> 82					•	
	gggccgggca	ccccagaag	ctgaccttga	gacaaggatt	tagatacaaa	60
	ggcaggtgcc					120
	cactacagge					180
	taggagatga					240
gaagcatggg	tatttatatg	tcagctccca	ttcattattg	gctgagggca	gctcctagag	300
ggcattgggt	ctgcgtttca	agcctgctgc	acataggctg	agaggaatcc	ctgagttcga	360
gtcacaggcg	cccacagtca	tgctcagaca	gcacatacag	gaacagtgac	tgcagggggc	420
ataggtggga	cacaaatacc	accagttata	aagaggaaag	atgggaagga	aagacaagag	480
	agttagattc	ctgggtcaga	tgtgaacccc	tggctctcaa	aacactcctt	540
ctttttttct						
				ctgttgcaca		600
	atcagggctc	gtggcagcct	ctacctccta	ggctcacatg	atcctcccac	660
ctcagcctcc	atcagggctc tgagtagctg	gtggcagcct ggactagagg	ctacctccta cacacatcac	ggctcacatg cacacttggc	atcctcccac tagtatttaa	660 720
ctcagcctcc atttttctgt	atcagggctc tgagtagctg agaagtccag	gtggcagcet ggactagagg gcgcagtggc	ctacctccta cacacatcac tcatgcctgt	ggctcacatg cacacttggc aatcccagca	atcctcccac tagtatttaa ctttgggagg	660 720 780
ctcagcctcc atttttctgt ccgaggcagg	atcagggctc tgagtagctg agaagtccag tggatcacct	gtggcagcct ggactagagg gcgcagtggc gaggtcagga	ctacctccta cacacatcac tcatgcctgt gttcaagacc	ggctcacatg cacacttggc aatcccagca agcctggcca	atcctcccac tagtatttaa ctttgggagg acatggtgaa	660 720 780 840
ctcagectee atttttetgt ccgaggeagg acccegecte	atcagggctc tgagtagctg agaagtccag tggatcacct tactaaaaat	gtggcagcct ggactagagg gcgcagtggc gaggtcagga acaaaaaaat	ctacctccta cacacatcac tcatgcctgt gttcaagacc tagcctggtg	ggctcacatg cacacttggc aatcccagca agcctggcca tcgtggcagg	atcctcccac tagtatttaa ctttgggagg acatggtgaa ctcctgtaat	660 720 780 840 900
ctcagcctcc atttttctgt ccgaggcagg accccgcctc cctggctcct	atcagggctc tgagtagctg agaagtccag tggatcacct tactaaaaat tgggaggctg	gtggcagcct ggactagagg gcgcagtggc gaggtcagga acaaaaaaat aggtaggaga	ctacctccta cacacatcac tcatgcctgt gttcaagacc tagcctggtg atcacttgta	ggctcacatg cacacttggc aatcccagca agcctggcca tcgtggcagg cccagaatgt	atcctcccac tagtatttaa ctttgggagg acatggtgaa ctcctgtaat ggagcttgca	660 720 780 840 900 960
ctcagcctcc atttttctgt ccgaggcagg accccgctc cctggctcct gtgagctgag	atcagggctc tgagtagctg agaagtccag tggatcacct tactaaaaat tgggaggctg atcatgccat	gtggcagcct ggactagagg gcgcagtggc gaggtcagga acaaaaaaat aggtaggaga tacactccag	ctacctccta cacacatcac tcatgcctgt gttcaagacc tagcctggtg atcacttgta	ggctcacatg cacacttggc aatcccagca agcctggcca tcgtggcagg cccagaatgt	atcctcccac tagtatttaa ctttgggagg acatggtgaa ctcctgtaat ggagcttgca	660 720 780 840 900 960
ctcagcctcc atttttctgt ccgaggcagg accccgctc cctggctcct gtgagctgag	atcagggctc tgagtagctg agaagtccag tggatcacct tactaaaaat tgggaggctg	gtggcagcct ggactagagg gcgcagtggc gaggtcagga acaaaaaaat aggtaggaga tacactccag	ctacctccta cacacatcac tcatgcctgt gttcaagacc tagcctggtg atcacttgta	ggctcacatg cacacttggc aatcccagca agcctggcca tcgtggcagg cccagaatgt	atcctcccac tagtatttaa ctttgggagg acatggtgaa ctcctgtaat ggagcttgca	660 720 780 840 900 960

<210> 83

<211> 1173

<212> DNA

48

.

```
<213> Homo sapiens
<220>
<221> SITE
<222> (548)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (603)
<223> n equals a,t,g, or c
<400> 83
gctgtctcag aaaaaagaaa aaagtttcta aagtaaaaat tgaaagtact tcccctacaa
                                                                       - 60
ccacaggttg ctttgacaga ttaatgtaaa ttcttccaga tactcttctg tggatgtaga
                                                                        120
aacatgcaga atgaggcaag ctttaatttg cttatgtcac ttactgtgga tagcctttca
                                                                        180
tatcttataa gttaatgtca gagcagcaat ctcattttt tccaatttgt aaacatttta
                                                                        240
tttaacctta tgatggatat tttggtggat ttcagtatta caaaaatgcc tattaatagt
                                                                       300
atattttcat tatatttctg ttacgaaatt ataatgctac aaacattact atgcctgtgg
                                                                       360
cagtatacat ctgcacaagt tttgaaaatg ttatgcattc ataggcaaaa atgggataac
                                                                        420
ttttgggcag tggtcatgat taatctgttg atcagaatcc agagattgcc cttctccttg
                                                                        480
ccaattgctt taagagtacm ctagtttttg gccgggtgca rtggctcatg cctgtatccc
                                                                        540
agcatttngg aggccaagac gggcggatca caaggtcagg agatcgggac catcctggct
                                                                        600
aanttggtga ggccccattc tactaaaaat tccaaaaaaa cccacmaaaa ccaaaaaaac
                                                                        660
ccrgccttgk tggtgggatt acaggcatgt gcmacaacac ccggctaatt tttttgtat
                                                                       720
ttttagtaga ggtggggtgt caccatgttg gccaggctga tctcaaactc ctggccttaa
                                                                       780
gtgatctgcc cacctcggac tcccaaagtg cggaattaca ggcgtgagcc accgcgcccg
                                                                       840
gccactggtt tttaaacttt attttgaaat tatttcaggc tgggcgcagt ggttcacgcc
                                                                       900
tgtgatccca acactttggg aggccgaggc gggcggatca cgaggtcagg agatcaagac
                                                                       960
catcctggct aaccccgtct ctactaaaaa tataaaaaat cagccgggca cggtggcagg
                                                                      1020
tgcctgtagt cccagctact cagtgggctg aggcaggaga atggtatgaa cccgggaggc
                                                                      1080
ggagcttgca gtgagctgag atcacgccac tgcactccag cctgggagac agagtgagac
                                                                      1140
tctgtctcaa aaaaaaaaa aaaaaaaactc gag
                                                                      1173
<210> 84
<211> 1561
<212> DNA
<213> Homo sapiens
<400> 84
ggcacgagtg aggctcatgt ctgacctgca gaactgtata atgataaatt atgttgtttg
                                                                        60
aaaccgctac atttgcggta atttgttaca gcagcaatag aaaacgaatc ccctgcccag
                                                                       120
aatgacttcc tcctttcctg tcggacgaag gctcaggcct tctgctgaaa gcttgctccc
                                                                       180
ctaagtagtc acacccaatg ccgaatactc cccagaagca gctgctattt tctgaggaca
                                                                       240
atgagttgct tgtaagcctg agaacaggac gaaaacccac tttgcaagca gccctgcgtg
                                                                       300
tgacgggcat gccctcggag ggcaggttgg ttttgctatc tgctttctgt cctgcctttt
                                                                       360
tecececatg ggteetgtet ggetettttg ettteteact ttgtgeagaa agecatetea
                                                                       420
actottetea caggagaata getgtatgga egtageaggg ggagttacea catgeetace
                                                                       480
tecatggttt tegagagggg ceeetgeeca aatgteteag tggeeacett cateagacea
                                                                       540
tggagcagtc agagcgggaa gggattctag agttggtcca gtccaaccat ctcatcttac
                                                                       600
atgtgaagga ggaaaggaag aaagggagaa aaataagaaa gctgaggtca accctcctac
                                                                       660
agggatgggc ctggccaaca ggatcccaag ggatgacata acattgaaat taagaaacca
                                                                       720
aggaaagttg agaactaaag aaaacagaac ccagtcagcc aagaggcatc cttgagggcc
                                                                       780
aaccaagccg tcaaacctgg atgcccccga cgagtcagaa agtcgggtgc ctcaagagcc
```

aagcagccaa	gaaatggggt	ctggaactgg	cactttggtc	cgcctctgtg	cactcaccca	900
gaaagggtgg	aagggaccct	gggaccaagt	gccaaggtca	cacaacggat	gaatagactg	960
				atcaccttcc		1020
aatgtttgtt	caaagccaca	aatgtatata	ctctttgtat	gtatacagat	tttttctaaa	1080
ggttaacatc	taaacagatc	aattaaggtc	agccttaatt	tgtctgagct	ttttggttaa	1140
				ctcctttctc		1200
				ggattacatt		1260
				tcatatagct		1320
				atactgtaca		1380
_				gaagagcatg	-	1440
-				gcgtttttca		1500
				agccaaaaaa		1560
a				_		1561
<210> 85						
<211> 1433						
<212> DNA						
<213> Homo	sapiens					
	-					
<400> 85						
cccggagccg	tggacgccct	acagctgaga	aggggaccca	aggggtcggc	cgcggccaag	60
				ttatkagaca		120
				gtggggagaa		180
		-		aactggccag		240
				gaccttggga		300
				tgacccacaa		360
				tcagcttccc		420
				tccatggcac		480
				aggctccgcc		540
				gtggagccct		600
				ttccccaact		660
				ctgcagacct		720
				cagctttata		780
				accctgaggg		840
				cggctagagg		900
				tccgaggctg		960
				gctcatatgg		1020
	-			gggcctcggg		1080
				gtggtcaggg		1140
				gcccttgact		1200
				gccctaggcc		1260
				cacagetetg		1320
				agtgagtgct		1380
				gggcccggta		1433
00000000			~~~~	3330003300		
<210> 86						
<211> 1377						
<212> DNA						
<213> Homo	sapiens					
<400> 86						
	ccagtectea	ttccatcttc	ttacaaotta	gggagctggg	tecaggeetg	60
				ggctccatgt		120
gaccaacyct	accucyaacc	-222-29	3anggett	ggccccacgc	gcaygt	120

PCT/US99/05804 WO 99/47540

50

cagggcaggt	cttcccccga	gtgatggctc	ttggactgtg	ctcctctggg	geceteteaa	180
ctctgtgtct	gtcatctgtc	acctgcctgg	ccattatggt	tttgatggca	gtggatgggc	240
tccatgggac	ttcaggcctg	gggtgagact	caggaccctg	gggtgggcat	ggatggggat	300
attggacccc	tgaaagaagg	gaagctgaga	gactttttc	ctttaaagac	ttttccatgt	360
tatctccact	cagagaattc	ttttctgcaa	agtcacggga	gggaggtgac	attgagccct	420
ccaatgtgac	agaaactgtg	ctgggaactt	tacatgtgtt	acctaatttg	tttaattatc	480
ccagcaactc	cacaaagtag	gcatttttat	tgttgaggaa	acagaagctt	agagactttg	540
tgagacttgc	ccgagacccc	aggtcacaca	ccagcaagga	tgaggtcaag	cttttaatcc	600
			tcacaacaat			660
			ttgactacat			720
			aaattttta			780
gttttctcta	gatatcggaa	gatgggctgt	gttttttgtc	tgttgtccag	tgaggctgat	840
ttgtagtcag	acaggtgagt	cagtttggtt	ggagtaggct	attgtggttc	tctctcatca	900
			tccaagttgg			960
attctccctt	aacaggctgt	gagggagtct	gggccaggca	ctaggccatg	agcagggcag	1020
actggggtaa	acccttagcg	agcctctctc	cagccacgag	gaaacctgga	gtgtgtgcgt	1080
gcctgtgtgc	tgctggtgtg	tgtgtgtgaa	tgcacacgtg	tgtgcatgca	ctgtgagctg	1140
			gtgtgtgtgt			1200
			tacatgtgta			1260
tgtttgtaca	gtgtgcgtgt	gtgtgtgtgt	ttgtgcacat	gagctgctgc	acacatataa	1320
gccttgtgaa	ttaggggaag	aagaaaggct	ccggcttaca	aaaaaaaaa	aaaaaaa	1377
010: 07						
<210> 87						
<211> 1715						
<212> DNA						
<213> Homo	sapiens					
<400> 87						
	acattggagc	tececacace	actcattgct	gcccaccagc	tatacaacta	60
			gaagccattg			120
			ggcatggcac			180
			tgatgtgtct			240
						2 3 0
	gaggagcaag	gagaggctga	gcggcacgtt		agttcttcgt	300
ggtgctcacc			gcggcacgtt caggctcact	ctgcggctac		
	agccagcgag	agctcttccc	caggctcact	ctgcggctac gctgacatgc	gccgcttccg	300
gaagccaccc	agccagcgag agactgcccc	agctcttccc ctgagccaga	caggctcact ggctcctggg	ctgcggctac gctgacatgc agttcagctg	gccgcttccg gcagccctgg	300 360
gaagccaccc ggaggcctca	agccagcgag agactgcccc gggcttattc	agctcttccc ctgagccaga tagcgcctgg	caggeteact ggeteetggg aceggeteet	ctgcggctac gctgacatgc agttcagctg ctgttcccac	geegetteeg geageeetgg caetggetge	300 360 420
gaagccaccc ggaggcctca agaggtgggc	agccagcgag agactgccc gggcttattc atggcacgag	agetetteee etgageeaga tagegeetgg caeggetgge	caggctcact ggctcctggg	ctgcggctac gctgacatgc agttcagctg ctgttcccac cggctggctg	gccgcttccg gcagccctgg cactggctgc gagggcactg	300 360 420 480
gaagccaccc ggaggcctca agaggtgggc ccgtcgggac	agccagcgag agactgccc gggcttattc atggcacgag accctttgga	agctcttccc ctgagccaga tagcgcctgg cacggctggc agcgcctctt	caggctcact ggctcctggg accggctcct tcagctggtg	ctgcggctac gctgacatgc agttcagctg ctgttcccac cggctggctg ccaccggggc	gccgcttccg gcagccctgg cactggctgc gagggcactg ctgatcgact	300 360 420 480 540
gaagccaccc ggaggcctca agaggtgggc ccgtcgggac gcggctaggg	agccagcgag agactgccc gggcttattc atggcacgag accctttgga gggcgcctgg	agetettece etgagecaga tagegeetgg caeggetgge agegeetett ecetggeaga	caggeteact ggeteetggg aceggeteet teagetggtg ettgetggag	ctgcggctac gctgacatgc agttcagctg ctgttcccac cggctggctg ccaccggggc ctcctagaag	gccgcttccg gcagccctgg cactggctgc gagggcactg ctgatcgact cagtccatgc	300 360 420 480 540
gaagccaccc ggaggcctca agaggtgggc ccgtcgggac gcggctaggg caaatccatt	agccagcgag agactgccc gggcttattc atggcacgag accctttgga gggcgcctgg ggggacatcg	agetettece etgagecaga tagegeetgg eaeggetgge agegeetett eeetggeaga acceceaget	caggeteact ggeteetggg aceggeteet teagetggtg ettgetggag getggaggaa ggaetgette	ctgcggctac gctgacatgc agttcagctg ctgttcccac cggctggctg ccaccggggc ctcctagaag ctatccatga	gccgcttccg gcagccctgg cactggctgc gagggcactg ctgatcgact cagtccatgc cggtctcctg	300 360 420 480 540 600
gaagccaccc ggaggcctca agaggtgggc ccgtcgggac gcggctaggg caaatccatt gtaccagagc	agccagcgag agactgccc gggcttattc atggcacgag accctttgga gggcgcctgg ggggacatcg ctgatcaaag	agetettece etgageeaga tagegeetgg eaeggetgge agegeetett eeetggeaga acceceaget ttetectaag	caggeteact ggeteetggg aceggeteet teagetggtg ettgetggag getggaggaa	ctgcggctac gctgacatgc agttcagctg ctgttcccac cggctggctg ccaccggggc ctcctagaag ctatccatga agagctgtcg	gccgcttccg gcagccctgg cactggctgc gagggcactg ctgatcgact cagtccatgc cggtctcctg ccatttccaa	300 360 420 480 540 600 660 720
gaagccaccc ggaggcctca agaggtgggc ccgtcgggac gcggctaggg caaatccatt gtaccagagc agcccagact	agccagcgag agactgccc gggcttattc atggcacgag accctttgga gggcgcctgg ggggacatcg ctgatcaaag tgggaactca	agetettece etgageeaga tagegeetgg eaeggetgge agegeetett eeetggeaga acceceaget tteteetaag gtacetggtt	caggeteact ggetectggg aceggeteet teagetggtg cttgetggag getggaggaa ggaetgette eegetteeee gegetgaate	ctgcggctac gctgacatgc agttcagctg ctgttcccac cggctggctg ccaccggggc ctcctagaag ctatccatga agagctgtcg agaagttcac	gccgcttccg gcagccctgg cactggctgc gagggcactg ctgatcgact cagtccatgc cggtctcctg ccatttccaa tgactgctct	300 360 420 480 540 600 660 720 780
gaagccaccc ggaggcctca agaggtgggc ccgtcgggac gcggctaggg caaatccatt gtaccagagc agcccagact gcgctagtgt	agccagcgag agactgccc gggcttattc atggcacgag accctttgga gggcgcctgg ggggacatcg ctgatcaaag tgggaactca tctggactcc	agetettece etgageeaga tagegeetgg eaeggeetett ecetggeaga acceceaget tteteetaag gtacetggtt acttaggaaa	caggeteact ggeteetgg aceggeteet teagetggtg ettgetggag getggaggaa ggaetgette eegetteeee gegetgaate gaegtetetg	ctgcggctac gctgacatgc agttcagctg ctgttcccac cggctggctg ccaccggggc ctcctagaag ctatccatga agagctgtcg agaagttcac aagtggtttt	gccgcttccg gcagccctgg cactggctgc gagggcactg ctgatcgact cagtccatgc cggtctcctg ccatttccaa tgactgctct ccgagaagcc	300 360 420 480 540 600 660 720 780 840
gaagccaccc ggaggcctca agaggtgggc ccgtcgggac gcggctaggg caaatccatt gtaccagagc agcccagact gcgctagtgt cttccagtac	agccagcgag agactgccc gggcttattc atggcacgag accctttgga gggcgcctgg ggggacatcg ctgatcaaag tgggaactca tctggactcc agccccagga	agetettece etgageeaga tagegeetgg eaeggetett ecetggeaga acceceaget ttetectaag gtacetggtt acttaggaaa eaegagaage	caggeteact ggeteetggg aceggeteet teagetggtg ettgetggag getggaggaa ggaetgette eegetteeee gegetgaate gaegtetetg eeeeetgeee	ctgcggctac gctgacatgc agttcagctg ctgttcccac cggctggctg ccaccggggc ctcctagaag ctatccatga agagctgtcg agaagttcac aagtggttt aactggtctc	gccgcttccg gcagccctgg cactggctgc gagggcactg ctgatcgact cagtccatgc cggtctcctg ccatttccaa tgactgctct ccgagaagcc cacctaccac	300 360 420 480 540 600 660 720 780 840 900
gaagccaccc ggaggcctca agaggtgggc ccgtcgggac gcggctaggg caaatccatt gtaccagagc agcccagact gcgctagtgt cttccagtac cacctggagt	agccagcgag agactgcccc gggcttattc atggcacgag accctttgga gggcgcctgg ggggacatcg ctgatcaaag tgggaactca tctggactcc agccccagga ctgtcatcaa	agetettece etgagecaga tagegeetgg eaeggeetett ecetggeaga acceceaget ttetectaag gtacetggtt acttaggaaa caegagaage caeageetgt	caggeteact ggeteetgg aceggeteet teagetggtg ettgetggag getggaggaa ggaetgette eegetteeee gegetgaate gaegtetetg	ctgcggctac gctgacatgc agttcagctg ctgttcccac cggctggctg ccaccggggc ctcctagaag ctatccatga agagctgtcg agaagttcac aagtggttt aactggtctc tggacccgcc	gccgcttccg gcagccctgg cactggctgc gagggcactg ctgatcgact cagtccatgc cggtctcctg ccatttccaa tgactgctct ccgagaagcc cactaccac tcctctgaag	300 360 420 480 540 600 660 720 780 840 900
gaagccaccc ggaggcctca agaggtgggc ccgtcgggac gcggctaggg caaatccatt gtaccagagc agcccagact gcgctagtgt cttccagtac cacctggagt ggagtggact	agccagcgag agactgcccc gggcttattc atggcacgag accctttgga gggcgcctgg ggggacatcg ctgatcaaag tgggaactca tctggactcc agccccagga ctgtcatcaa ggaccactga	agetettece etgageeaga tagegeetgge agegeetett ecetggeaga acceeaget ttetectaag gtacetggtt acttaggaaa caegagaage caeageetgt atgteactgt	caggeteact ggetectggg aceggeteet teagetggtg ettgetggag getggaggaa ggaetgette eegetteeee gegetgaate gaegtetetg eeeeetteee teaeeette	ctgcggctac gctgacatgc agttcagctg ctgttcccac cggctggctg ccaccggggc ctcctagaag ctatccatga agagctgtcg agaagttcac aagtggttt aactggtctc tggacccgcc atgggcctac	gccgcttccg gcagccctgg cactggctgc gagggcactg ctgatcgact cagtccatgc cggtctcctg ccatttccaa tgactgctct ccgagaagcc cacttaccac tcctctgaag cagattgcct	300 360 420 480 540 600 660 720 780 840 900 960
gaagccaccc ggaggcctca agaggtgggc ccgtcgggac gcggctaggg caaatccatt gtaccagagc agcccagact gcgctagtgt cttccagtac cacctggagt ggagtggact gccagaggca	agccagcgag agactgcccc gggcttattc atggcacgag accctttgga gggcgcctgg ggggacatcg ctgatcaaag tgggaactca tctggactcc agccccagga ctgtcatcaa ggaccactga ggactgacca	agetettece etgageeaga tagegeetgg eaeggeetett eeetggeaga acceceaget tteteetaag gtacetggtt acttaggaaa eaegagaage caeageetgt atgteaetgt geeettetgg	caggeteact ggetectggg aceggeteet teagetggtg ettgetggag getggaggaa ggaetgette eegetteeee gegetgaate gaegtetetg eeeetgeee tteaeeette teettgaate	ctgcggctac gctgacatgc agttcagctg ctgttcccac cggctggctg ccaccggggc ctcctagaag ctatccatga agagctgtcg agaagttcac aagtggttt aactggtctc tggacccgcc atgggcctac aagccagaca	gccgcttccg gcagccctgg cactggctgc gagggcactg ctgatcgact cagtccatgc cggtctcctg ccatttccaa tgactgctct ccgagaagcc cacctaccac tcctctgaag cagattgcct ctgagtgaca	300 360 420 480 540 600 660 720 780 840 900 960 1020 1080
gaagccaccc ggaggcctca agaggtgggc ccgtcgggac gcggctaggg caaatccatt gtaccagagc agcccagact gcgctagtgt cttccagtac cacctggagt ggagtggact gccagaggca ccaaaggctt	agccagcgag agactgcccc gggcttattc atggcacgag accctttgga gggggcctgg ggggacatcg ctgatcaaag tgggaactca tctggactcc agccccagga ctgtcatcaa ggaccactga ggactgacca tgtaactatg	agetettece etgageeaga tagegeetgg eaeggetett ecetggeaga acceeaget ttetectaag gtacetggtt acttaggaaa caegagaage caeageetgt atgteaetgt geeettetgg tettgagggt	caggeteact ggetectggg aceggeteet teagetggtg ettgetggag getggaggaa ggaetgette eegetteeee gegetgaate gaegtetetg eeceetgeee tteaecette teettgaate geeceaggge	ctgcggctac gctgacatgc agttcagctg ctgttcccac cggctggctg ccaccggggc ctcctagaag ctatccatga agagctgtcg agaagttcac aagtggttt aactggtctc tggacccgcc atgggcctac aagccagaca agcctggcag	gccgcttccg gcagccctgg cactggctgc gagggcactg ctgatcgact cagtccatgc cggtctcctg ccatttccaa tgactgctct ccgagaagcc cactaccac tcctctgaag cagattgcct ctgagtgaca caggaaccgc	300 360 420 480 540 600 660 720 780 840 900 960 1020 1080
gaagccaccc ggaggcctca agaggtgggc ccgtcgggac gcggctaggg caaatccatt gtaccagagc agcccagact gcgctagtgt cttccagtac cacctggagt ggagtggact gccagaggca ccaaaggctt cctccccaaa	agccagcgag agactgcccc gggcttattc atggcacgag accctttgga gggggacatcg ctgatcaaag tgggaactca tctggactcc agccccagga ctgtcatcaa ggaccactga ggactgacca tgtaactatg cacccacagc	agetettece ctgagecaga tagegectgg caeggetett ceetggeaga acceeaget ttetectaag gtacetggtt acttaggaaa caegagaage caeageetgt atgteaetgt geeettetgg tettgagggt caetgaecea	caggeteact ggetectggg aceggeteet teagetggtg ettgetggag getggaggaa ggaetgette eegetteeee gegetgaate gaegtetetg eeceetgeee tteaeeette teettgaate geeceaggge etgetgeee	ctgcggctac gctgacatgc agttcagctg ctgttcccac cggctggctg ccaccggggc ctctagaag ctatccatga agagctgtcg agaagttcac aagtggttt aactggtctc tggacccgcc atgggcctac aagccagaca agcctggcag cagagagtca	gccgcttccg gcagccctgg cactggctgc gagggcactg ctgatcgact cagtccatgc cggtctcctg ccattcccaa tgactgctct ccgagaagcc cactaccac tcctctgaag cagattgcct ctgagtgaca caggaaccgc ggtcaacccc	300 360 420 480 540 600 660 720 780 840 900 960 1020 1080 1140
gaagccaccc ggaggcctca agaggtgggc ccgtcgggac gcggctaggg caaatccatt gtaccagagc agcccagact gcgctagtgt cttccagtac cacctggagt ggagtggact gccagaggca ccaaaggctt cctccccaaa gaggaccct	agccagcgag agactgcccc gggcttattc atggcacgag accctttgga ggggacatcg ctgatcaaag tgggaactca tctggactcc agccccagga ctgtcatcaa ggaccactga ggactgacca tgtaactatg cacccacagc tgggcccttc	agetettece ctgagecaga tagegectegg caeggetett ceetggeaga acceeaget ttetectaag gtacetggtt acttaggaaa caegagaage caeageetgt atgteaetgt geettetgg tettgagggt caetgaecea tggggtaete	caggeteact ggetectggg aceggeteet teagetggtg ettgetggag getggaggaa ggaetgette eegetteee gegetgaate gaegtetetg eeeeettet teaeeette teettgaate geeeeagge etgetgeee teeaggaete	ctgcggctac gctgacatgc agttcagctg ctgttcccac cggctggctg ccaccggggc ctctagaag ctatccatga agagctgtcg agaagttcac aagtggttt aactggtctc tggacccgcc atgggcctac aagccagaca agcctggcag cagagagtca ccctggtaga	gccgcttccg gcagccctgg cactggctgc gagggcactg ctgatcgact cagtccatgc cggtctcctg ccatttccta tgactgctct ccgagaagcc cactaccac tcctctgaag cagattgcct ctgagtgaca caggaaccgc ggtcaacccc gtctcggag	300 360 420 480 540 600 660 720 780 840 900 960 1020 1080 1140 1200 1260
gaagccaccc ggaggcctca agaggtgggc ccgtcgggac gcggctaggg caaatccatt gtaccagagc agcccagact gcgctagtgt cttccagtac cacctggagt ggagtggact gccagaggca ccaaaggctt cctcccaaa gaggacccct ttcacacagg	agccagcgag agactgcccc gggcttattc atggcacgag accctttgga gggggacatcg ctgatcaaag tgggaactca tctggactcc agccccagga ctgtcatcaa ggaccactga ggactgacca tgtaactatg cacccacagc tgggcccttc gtggcaaaca	agetettece ctgagecaga tagegeetgg caeggetett ceetggeaga acceeaget tteteetaag gtacetggtt acttaggaaa caegagaage caeageetgt atgteaetgt geeettetgg tettgagggt caetgaecea tggggtaete eeeectagag	caggeteact ggetectggg aceggeteet teagetggtg ettgetggag getggaggaa ggaetgette eegetteee gegetgaate gaegtetetg eeceetgeee tteaeeette teettgaate geeeagge etgetgeee teeaggaete ettteggeee	ctgcggctac gctgacatgc agttcagctg ctgttcccac cggctggctg ccaccggggc ctcctagaag ctatccatga agagctgtcg agaagttcac aagtggttt aactggtctc tggacccgcc atgggcctac aagccagaca agcctggcag cagagagtca ccctggtaga tgaatcctgc	gccgcttccg gcagccctgg cactggctgc gagggcactg ctgatcgact cagtccatgc cggtctcctg ccatttccaa tgactgctct ccgagaagcc cactaccac tcctctgaag cagattgctt ctgagtgaca caggaaccgc ggtcaacccc gtctcggag ccctagcct	300 360 420 480 540 600 660 720 780 840 900 960 1020 1080 1140 1200 1260 1320
gaagccaccc ggaggcctca agaggtgggc ccgtcgggac gcggctaggg caaatccatt gtaccagact gcgctagtgt cttccagtac cacctggagt ggagtggact gccagaggca ccaaaggctt cctccccaaa gaggacccct ttcacacagg	agccagcgag agactgcccc gggcttattc atggcacgag accctttgga ggggacatcg ctgatcaaag tgggaactca tctggactcc agccccagga ctgtcatcaa ggaccactga ggactgacca tgtaactatg cacccacagc tgggcccttc gtggcaaaca tcagccacct	agetettece ctgagecaga tagegectegg caeggetett cectggeaga acceceaget ttetectaag gtacetggtt acttaggaaa caegagaage caeageetgt atgteactgt gecettetgg tettgagggt caetgaecea tggggtaete cecectagag gtgtecettg	caggeteact ggetectggg aceggeteet teagetggtg ettgetggag getggaggaa ggaetgette eegetteeee gegetgaate gaegtetetg eeceetgeee teaeeette teettgaate geeceaggge etgetgeee teeaggaete etteegeee	ctgcggctac gctgacatgc agttcagctg ctgttcccac cggctggctg ccaccggggc ctcctagaag ctatccagaag agagctgtcg agaagttcac aagtggttt aactggtctc tggacccgcc atgggcctac aagccagaca agcctggcag cagagagtca cctggtaga tgaatcctgc tcttcacttc	gccgcttccg gcagccctgg cactggctgc gagggcactg ctgatcgact cagtccatgc cggtctcctg ccatttccaa tgactgctct ccgagaagcc cactaccac tcctctgaag cagattgctc ctgagtgaca caggaaccgc ggtcaacccc gtctcggag ccctagcag ccctagcct ccacttggac	300 360 420 480 540 600 660 720 780 840 900 960 1020 1080 1140 1200 1260 1320 1380

aagcacttgc ggccaccgtc aagtcccttt gctctcggac cctgggtttc tcatccttta

```
atgaggtggg ttcagaagct ctcccatctt cacagcaacc ctggcactgg cttctcaatg
                                                                     1620
                                                                     1680
ggagggaagt cagcagagaa actgaagtgt tagacactat gtgtcccacc accccattac
                                                                     1715
agagacatat gacaatgaaa aaaaaaaaaa aaaaa
<210> 88
<211> 417
<212> DNA
<213> Homo sapiens
<400> 88
                                                                       60
ccacgcgtcc gctcctctag aggctccaca tgaagtccca gtgctacagt cctagttatt
ttgccttctt ctgcctggtt ttctttcaga tcacctcagc cagttctcag acacttaggg
                                                                      120
gacatgttct ctgcaggacc actctgaggg actcttctgc atattgctga cctgagagga
                                                                      180
                                                                      240
tggcctcaga gctgacttgg gcaatcctcc ccaacaggaa ggggagacat tgcctgccac
                                                                      300
tgaggaaaca ggtcatgaag gtggagataa gctgcaaggg gcgaagcaac tttatgtcag
                                                                      360
tggaaaacgt gtctctttaa agctgctatg tgaacagctt ttacagtcat taaatttacc
                                                                      417
<210> 89
<211> 1167
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (432)
<223> n equals a,t,g, or c
<400> 89
                                                                       60
gggggtgggg caggcgacgg tggggaagat ggcgtaccag agcttgcggc tggagtacct
                                                                      120
gcagatecca ceggteagee gegeetacae caetgeetge gteeteacca cegeegeegt
                                                                      180
gcagttggaa ttgatcacac cttttcagtt gtacttcaat cctgaattaa tctttaaaca
ctttcaaata tggagattaa tcaccaactt cttatttttt gggccagttg gattcaattt
                                                                      240
                                                                      300
tttatttaac atgatttttc tatatcgtta ctgtcgaatg ctagaagaag gctctttccg
aggteggaea geagaetttg tatttatgtt eetttttggt ggattettaa tgaeeetttt
                                                                      360
tggtctgttt gtgagcttag ttttcttggg ccaggccttt acaataatgc tcgtctatgt
                                                                      420
gtggagccga angaacccct atgtccgcat gaacttcttc ggccttctca acttccaggc
                                                                      480
cccctttctg ccctgggtgc tcatgggatt ttccttgttg ttggggaact caatcattgt
                                                                      540
ggaccttttg ggtattgcag ttggacacat atatttttc ttggaagatg tatttcccaa
                                                                      600
                                                                      660
tcaacctggt ggaataagaa ttctgaaaac accatctatt ttgaaagcta tttttgatac
                                                                      720
accagatgag gatccaaatt acaatccact acctgaggaa cggccaggag gcttcgcctg
                                                                      780
gggtgagggc cagcggcttg gaggttaaag cagcagtgcc aataatgaga cccagctggg
aaggactcgg tgatacccac tgggatcttt tatcctttgt tgcaaaagtg tggacacttt
                                                                      840
                                                                      900
 tgacagettg geagatttta acteeagaag caetttatga aatggtacae tgactaatee
                                                                      960
 agaagacatt tccaacagtt tgccagtggt tcctcactac actggtactg aaagtgtaat
ttcttagagc caraaaactg gagaaacaaa tatcctgcca cctctaacaa gtacatgagt
                                                                      1020
 acttgatttt tatggtataa gcagagcctt ttcttcctct tcttgataga tgaggccatg
                                                                      1080
 gtgtaaatgg aagtttcaga gaggacaaaa taaaacggaa ttccattttt ctctcactgt
                                                                      1140
                                                                      1167
 aaaaaaaaa aaaaaaaggg cggccgc
```

<210> 90 <211> 1892

```
<212> DNA
<213> Homo sapiens
<400> 90
                                                                        60
ccacgcgtcc gcgggaccgg acggatcttc tccggccatg aggaagccag ccgctggctt
                                                                       120
cettecetea etectgaagg tgetgeteet geetetggea eetgeegeag eeeaggatte
                                                                       180
gactcaggcc tccactccag gcagccctct ctctcctacc gaatacgaac gcttcttcgc
                                                                       240
actgctgact ccaacctgga aggcagagac tacctgccgt ctccgtgcaa cccacggctg
coggaatoco acactogtoo agotggacca atatgaaaac cacggottag tgocogatgg
                                                                       300
                                                                       360
tgctgtctgc tccaacctcc cttatgcctc ctggtttgag tctttctgcc agttcactca
                                                                       420
ctaccgttgc tccaaccacg tctactatgc caagagagtc ctgtgttccc agccagtctc
tattetetea ectaacaete teaaggagat agaagettea getgaagtet caeceaceae
                                                                       480
                                                                       540
gatgacetec eccateteac eccaetteac agtgacagaa egecagacet tecagecetg
                                                                       600
gcctgagagg ctcagcaaca acgtggaaga gctcctacaa tcctccttgt ccctgggaag
ccaggagcaa gcgccagagc acaagcagga gcaaggagtg gagcacaggc aggagccgac
                                                                       660
                                                                       720
acaagaacac aagcaggaag aggggcagaa acaggaagag caagaagagg aacaggaaga
                                                                       780
ggagggaaag caggaagaag gacaggggac taaggaggga cgggaggctg tgtctcagct
gcagacagac tcagagccca agtttcactc tgaatctcta tcttctaacc cttcctcttt
                                                                       840
tgctccccgg gtacgagaag tagagtctac tcctatgata atggagaaca tccaggagct
                                                                       900
                                                                       960
cattegatea geccaggaaa tagatgaaat gaatgaaata tatgatgaga acteetaetg
gagaaaccaa aaccctggca gcctcctgca gctgccccac acagagcctt gctggtgctg
                                                                      1020
tgctattcga tcgtggagaa tacctgcatc ataaccccca cagccaaggc ctggaagtac
                                                                      1080
                                                                      1140
atggaggagg agateettgg tttegggaag teggtetgtg acageettgg geggegaeae
                                                                      1200
atgtctacct gtgccctctg tgacttctgc tccttgaagc tggagcagtg ccactcagag
gccagcctgc ageggcaaca atgcgacacc teccacaaga etecetttge ageccettge
                                                                      1260
                                                                       1320
ttgcctccca gagcctgtcc atcggcaacc aggtagggtc cccagaatca ggccgctttt
acgggctgga tttgtacggt gggctccaca tggacttctg gtgtgcccgg cttgccacga
                                                                       1380
aaggetgtga agatgteega gtetetgggt ggeteeagae tgagtteett agetteeagg
                                                                       1440
atggggattc cctaccaaga tttgtgacac agactatatc cagtacccaa actactgttc
                                                                       1500
                                                                       1560
cttcaaaagc cagcagtgtc tgatgagaaa ccgcaatcgg aaggtgtccc gcatgagatg
tetgeagaat gagaettaca gtgegetgag ceetggeaaa agtgaggaeg ttgtgetteg
                                                                       1620
                                                                       1680
atggagccag gagttcagca ccttgactct aggccagttc ggatgagctg gcgtctattc
tgcccacacc ccagcccaac ctgcccacgt tctctattgt tttgagaccc cattgctttc
                                                                       1740
aggetgeece ttetgggtet gttactegge ecetacteae attteettgg gttggageaa
                                                                       1800
cagtcccaga gagggccatg gtgggagtgc gccctcctta aaagatgact ttacataaaa
                                                                       1860
                                                                       1892
tgttgatctt caaaaaaaaa aaaaaaaaaa aa
<210> 91
<211> 523
<212> DNA
<213> Homo sapiens
<400> 91
cacagcaaag caagttctaa gagccaagct tcagaccaat cccccaccgt gaagtccccc
                                                                         60
                                                                        120
gtgcgagtgc cccttgaagg aggttttcta acaggtgagt ggtctgattc tgtctctgtc
ctgtgggatg gatgggctgg cacttgatgg ctctccttcc ccctcacccy ccacggagaa
                                                                        180
ggctggaagg tgcatttctc agacttcctt gcctgggaaa tgggaagtga tgcagaggat
                                                                        240
cccaacgtct cctcggcagg cttggtggtg gacgtgctgg gccatgttcc aggggccagc
                                                                        300
                                                                        360
tgctggctcc gtgggtgctg agaggaaggg ggaaggctgt ctattttttg gccaggatga
 atccagcaga tgtggtaggt cctggccgct tgctgacccg tgggtctacc gggtgctccg
                                                                        420
                                                                        480
 gagetaatgg tececagatg etceacegte etgatgtgge agaggeatgg cattitggeg
                                                                        523
 ggccagtttg gtggcatcct gggaaccgtt ttggaggctc gag
```

```
<210> 92
<211> 1382
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (1382)
<223> n equals a,t,g, or c
<400> 92
                                                                       60
gccggctggc agcacgactc gcgtagccgt gcgccgattg cctctcggcc tgggcaatgg
                                                                      120
tcccggctgc cggtcgacga ccgcccgcg tcatgcggct cctcggctgg tggcaagtat
                                                                      180
tgctgtgggt gctgggactt cccgtccgcg gcgtggaggt tgcagaggaa agtggtcgct
                                                                      240
tatggtcaga ggagcagcct gctcaccctc tccaggtggg ggctgtgtac ctgggtgagg
                                                                      300
aggageteet geatgaceeg atgggeeagg acagggeage agaagaggee aatgeggtge
                                                                      360
tggggctgga cacccaaggc gatcacatgg tgatgctgtc tgtgattcct ggggaagctg
                                                                      420
aggacaaagt gagttcagag cctagcggcg tcacctgtgg tgctggagga gcggaggact
                                                                      480
caaggtgcaa cgtccgagag agccttttct ctctggatgg cgctggagca cacttccctg
                                                                      540
acagagaaga ggagtattac acagagccag aagtggcgga atctgacgca gccccgacag
                                                                      600
aggactccaa taacactgaa agtctgaaat ccccaaaggt gaactgtgag gagagaaaca
ttacaggatt agaaaatttc actctgaaaa ttttaaatat gtcacaggac cttatggatt
                                                                      660
                                                                      720
ttctgaaccc aaacggtagt gactgtactc tagtcctgtt ttacaccccg tggtgccgct
                                                                      780
tttctgccag tttggcccct cactttaact ctctgccccg ggcatttcca gctcttcact
                                                                      840
ttttggcact ggatgcatct cagcacagca gcctttctac caggtttggc accgtagctg
                                                                      900
ttcctaatat tttattattt caaggagcta aaccaatggc cagatttaat catacagatc
                                                                      960
gaacactgga aacactgaaa atcttcattt ttaatcagac aggtatagaa gccaagaaga
                                                                     1020
atgtggtggt aactcaagcc gaccaaatag gccctcttcc cagcactttg ataaaaagtg
                                                                     1080
tggactggtt gcttgtattt tccttattct ttttaattag ttttattatg tatgctacca
                                                                     1140
ttcgaactga gagtattcgg tggctaattc caggacaaga gcaggaacat gtggagtagt
                                                                     1200
gatggtctga aagaagttgg aaagaggaac ttcaatcctt cgtttcagaa attagtgcta
                                                                     1260
cagtttcata cattttctcc agtgacgtgt tgacttgaaa cttcaggcag attaaaagaa
                                                                     1320
tcatttgttg aacaactgaa tgtataaaaa aattataaac tggtgtttta actagtattg
                                                                     1380
1382
an
<210> 93
<211> 1747
<212> DNA
<213> Homo sapiens
<400> 93
                                                                       60
ccacgcgtcc ggctacctgt gcatcgtgct gctcatgctg ctgctgctca tcttctggat
                                                                      120
cgcgccggcc catgggccca ccaacatcat ggtctacatc agcatctgct ccttgctggg
                                                                      180
cagtttcacc gtgccttcca ccaagggcat cgggctggcg gcccaagaca tcttgcataa
                                                                      240
caaccegtce agtcagagag ccctctgcct gtgcctggta ctcctggccg tgctcggctg
                                                                      300
cagcatcatc gtccagttca ggtacatcaa caaggcgctg gagtgcttcg actcctcggt
                                                                      360
gttcggggcc atctactacg tcgtgtttac cacgctggtc ctgctggcct cagccatcct
                                                                      420
cttccgggag tggagcaacg tgggcctggt ggacttcttg gggatggcct gtggattcac
                                                                      480
gaccgtctcc gtggggattg tccttataca ggtgttcaaa gagttcaatt tcaaccttgg
ggagatgaac aaatctaata tgaaaacaga ctagattgca ataggagctt ggatggttcg
                                                                      540
aggaataggc attggaggtg gtttctggcc gtgattggat gtgaagtaga agaggtcctc
                                                                      600
gatcatggtg ttagaattga ctggatagta acaggtggtc tggtggatag cggggagcat
                                                                      660
ggctcagcac cagagcagag gcccagcagc ctctgcagcc caaacgtccc aacggtgcct
                                                                      720
```

PCT/US99/05804

780

840

WO 99/47540

```
54
```

ggaccatctc ttctgatgag acgaatctca ttttcatttc cattaacctg gaagctttca.

tgaatatttc ttctttaaaa cattttaaca ttatttaaac agaaaaagat gggctctttc

```
900
tgggtaggtg gtacatgata gcagagatat ttttacttag attactttgg gaatgagaga
                                                                     960
ttgtgtcttg aactctgcac tgtacaggat gtgtctgtag ttgtgttagt ttgcattaag
catgtataca ttcaagtatg tcatccaaat aagaggcata tcattgaatt gtttttaatc
                                                                    1020
ctctgacaag ttgactcttc gaccccacc cccacccaag acattttaat agtaaataga
                                                                    1080
gagagagaga agagttaatg aacatgaggt agtgttccac tggcaggatg acttttcaat
                                                                    1140
agctcaaatc aatttcagtg cctttatcac ttgaattatt aacttaattt gactcttaat
                                                                    1200
qtqtatatgt tcttagatta gaataatgca acttcgagta tgctttaata tttcaatatt
                                                                    1260
                                                                    1320
caagttacaa atgtataagg cagttagaaa taatacagtc acatgtcact taatgatagg
                                                                    1380
gaaacattct gagaaatgca ttgtaaggtg actttattgt gtgaacatca tggagtgcac
ttatacaaac ctagatggga cacctatgac ccacccaggc cagatggtac agcctgttgc
                                                                    1440
tectgggeca cacacetgta cagcatgtga etgeactgaa tacegeagge aattgtaaca
                                                                    1500
cagtggtgag tatttgtgtt tacaaacata ggaaaggtac agtaaaacta tggtattaca
                                                                    1560
atgttatggg accaccgtca tgtaagtggt atgtctttga cagaaacatg gttacgtggt
                                                                    1620
tcatgactgt atattcactg gaagatagtc aagactaaag acacattaga gcaaattgac
                                                                    1680
1740
                                                                    1747
aaaaaaa
<210> 94
<211> 600
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (553)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (560)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (589)
<223> n equals a,t,g, or c
                                                                      60
gaatteggea egageggeae gageegagat egttetgggg etgetggtat ggaegettat
tgctggaact gagtacttcc gggtccccgc atttggctgg gtcatgtttg tagctgtatt
                                                                     120
ttactgggtc ctcaccgtst tcttcctcat tatctacata acaatgacct acaccaggat
                                                                     180
tccccaggtg ccctggacaa cagtgggcct gtgctttaac ggcagtgcct tcgtcttgta
                                                                     240
cctctctgcc gctgttgtag atgcatcttc cgtctcccct gagaaggaca gtcacaactt
                                                                     300
                                                                     360
caacagctgg gcggcctcat cgttctttgc cttcctggtc accatctgct acgctggaaa
                                                                     420
tacatatttc agttttawag catggagawc caggaccata cagtgattta ccattttgat
                                                                     480
aattaaaagg aaaaaaaag gaagactctc actgtaaaaa cagctgtagg tataatgtat
attcccagag aattgtattt aactaattaa tgttttttat attcttaaat ttgctcacaa
                                                                     540
attgtggttt gtnacaattn aactgggtta ctttatttgg caagtgttnt aggcttttaa
                                                                     600
```

<210> 95

<211> 586

```
<212> DNA
<213> Homo sapiens
<400> 95
                                                                     60
ggcacgaggt tttttccttt ataacggaag ttttataatt catcttttat gtaagtgtaa
                                                                    120
ttctcattaa aaatacccta aagcttaaag tttgcaaggc tgcccagcct aacccacaac
agtttgatgc tgccccctag cgtttgattc ccttcacctt ttgctaaaat aaggtaatgt
                                                                    180
                                                                    240
ttaaattaca attagattta cttactgctg taaatctggt ctattttagt ttcctctggg
                                                                    300
tagttagtgt tgctaataag atggacgtaa gtgtttttga actggtgaat tctgattgct
                                                                    360
tttagccccc agttttccaa ataggggtga attttgggta gagatagaac aatcaccaag
ttaccttgct ccaaaaaaga aatttacgta tgggattgtt ttcaaagcgg gaagttagct
                                                                    420
gtgtaaataa caacaatttt atatatttaa tctgggcttc tccttatctt gaatgatata
                                                                     480
aaaatctact ttctagatta atttagttcc atataacttt gtattgcttt gactgtactg
                                                                    540
586
<210> 96
<211> 802
<212> DNA
<213> Homo sapiens
<400> 96
                                                                     60
ggcacgagcc ctcctccctg ctcgccccca gattcccctc ccctccctgg tgcttttgtc
                                                                     120
tggagggtgt tatgggtttg tgtgtgtatg agcgtgtgtg tgtttttgga tttcagacta
                                                                     180
attttctgga gtttctgccc ctgctctgcg tcaccctcac gtcacttcgc cagcagtagc
                                                                     240
agaggeggeg geggeggete eeggaattgg gttggageag gageeteget ggetgetteg
                                                                     300
ctcgcgctct acgcgctcag tccccggcgg tagcaggagc ctggacccag gcgccgccgg
                                                                     360
cgggcgtgag gcgccggagc ccgggtgagc agcgcagata gtgccctcgg tcgcctcggc
                                                                     420
cctcactgtc tccccctggg gcggcctcgg ctactcccca ggtgggacgt gccgcgccac
                                                                     480
ctgcccgcgc caccggcacc cagcggccgt ggcggattct gcagcatcat tcgggggccc
                                                                     540
cgtcgcggag ccaaagccgc cggcagtctc cgcattcccc tttaaagggt ccttcgcccg
                                                                     600
660
ggctcgaaga gagagtgggc acactggtgg ctccagcggc gtcagtgcca tcgcggggca
                                                                     720
agttgattcc tgggcactca tccatccaca gtctccgggc tgggggtcggg gtggggatga
                                                                     780
cgcgagcaga gagggagagt gccccaatta gtggtgttgg gggtcctacg ctcagtctta
                                                                     802
cgcgtgtctg tttgtcctca gc
<210> 97
<211> 1226
<212> DNA
<213> Homo sapiens
<400> 97
ggcacgagca tgctttgctt acaatggagt ctgcagtgag gggagatgct gggatagcca
                                                                      60
tttccatggc tctgttatgc aagcacaaat ttcatctcct agatggactt cctggttttc
                                                                     120
                                                                     180
tettactgca gtaacactgg cettecette tetaatteet taccecaget geggeatece
                                                                     240
tgtgttaact caggatgcca agtggccctc agattacact tctccagata gctgaatgag
                                                                     300
tetgetttea etgtgaetgg gacetgaatg acetgeagte agggeeeaga gttgggaete
                                                                     360
tatactaccc tgggctctgg tctgtaggtt tgtagtagcc accggtaata agccaagggc
                                                                     420
taggetettg tttgagttta tggecacetg gaattttcag teateteatg atacaggegg
                                                                     480
gaggggcaga acagatagat tacgacaggt ttggttttta aattttccaa ccaagtggaa
aggcaagttg gtcttataga aagcactact gcacttagta gctatgtgat tttgagcaaa
                                                                     540
ccacataatc tctctaggtc cattttccta accacaagat aaagatgtta cattgtcaaa
                                                                     600
gettgeegta gatttggggt gaatgaaaat tatteettge ttteateact acetttatag
                                                                     660
```

ctctcatcac	tacctttata	gctcatcact	gtgccttttt	ttctttccta	agaaagacat	720
	tcctctcctc					780
	agatgaagct					840
	aagcctgact					900
	acaatgccta					960
	tcagacattt					1020
	cttttattga					1080
	atgtgcaatg	-			=	1140
	attttagtgt				~	1200
-	aaaaaaaaa					1226
-						
<210> 98	•					
<211> 1120						
<212> DNA						
<213> Homo	sapiens					
<400> 98						
aggggactct	caccctctcc	cagcaatgtc	taaagtcagg	catctgaaaa	ccagcagtaa	60
tectgeetet	gaagtttatc	aggaaaggag	cttaaaagag	aaccaaattc	agcctgtgtt	120
ggaactctca	gtcccagagg	ggtgtggttt	atagctctcc	ggcctgctgt	tggacttagg	180
ctgtgaccca	cagaaggacg	ccagaaagta	ctcaagacat	tcacggtgcc	ccggtcagca	240
ctcgccatga	cgaagacttc	tacatgcata	taccacttcc	ttgttctgag	ctggtatact	300
ttcctcaatt	attacatctc	acaggaagga	aaagacgagg	tgaaacccaa	aatcttggca	360
aatggtgcaa	ggtggaaata	tatgacgctg	cttaatctgc	tcttgcagac	cattttctac	420
ggggtcacct	gcctggatga	tgtgctgaaa	agaaccaaag	ggggaaaaga	cattaagttc	480
ctaactgcct	tcagagacct	gcttttcacc	actctggctt	ttcctgtatc	cacgtttgta	540
tttttggcat	tctggatcct	ctttctctac	aatcgagatc	tcatttaccc	caaggtccta	600
-	tccccgtgtg					660
ttggctgaag	tcgtcctcag	gcctcactcc	tatccatcaa	agaagacagg	actcaccttg	720
stggctgctg	ccagcattgc	ttacatcagc	cgcatcctat	ggctctactt	tgagacgggt	780
	atcctgtgtt					840
ctcagctacg	tcttcatcgc	cagcatctac	ctacttggag	agaagctcaa	ccactggaaa	900
	tgaggcagcc			_		960
ccaagaaaga	agaaaacaca	agagattttt	ctcatctttt	tttttttt	tctggtggag	1020
ggaggtggtg	gaggaacata	gcaaagtagg	agggacagag	agtgatactt	aaatttaata	1080
agaggttcgt	gaaggtaaaa	aaaaaaaaa	aaaactcgag			1120
	•					
<210> 99						
<211> 2596						
<212> DNA						
<213> Homo	sapiens					
400 00						
<400> 99						
	gacttggcaa		· -	_		60
_	gatggatttc					120
	tggtctgatg	-	_		_	180
	tgtgttctct					240
	ggtgctattt	_		• •	_	300
	agatgtccag					360
_	aagtctgatc		-			420
	cactctgcaa					480
	atgagggaaa			gtttggcttg		540 600

ggaaaataaa	agtaagagtg	gtgttggaag	atgtagcatg	ttgatagctt	caaatgaaac	660
	gttgaatgtg					720
	acagcaatag					780
	tggttcctct					840
	caaggagtga					900
	gttttctaat					960
	gtatttactc					1020
	catgtgctgg					1080
	agaatggagg					1140
	taatgttcat					1200
	tttttaacaa					1260
	tttaattttt					1320
	tgtaaactct					1380
	ataagacatg					1440
	aggctggagt					1500
	cgattctcct					1560
	gctaattttt					1620
	actcgtgacc					1680
	gagccaccgc					1740
	atgtttctta					. 1800
tgaagagttt	ttttctgttt	ataattgttt	ataaaacaat	gacttttaaa	gatttagtgc	1860
	cccaagtgtg					1920
	ttttctgatg					1980
	tgataatttt					2040
	ttacttttt					2100
	atttgaaaaa					2160
	taaaattatg					2220
	tagattatag					2280
attcagcatt	taacttacat	ttgtacttta	gagtatttt	gtataaaatc	catagattta	2340
	agagtattta					2400
	acagttgtcc					2460
	tacaatttta					2520
tgaaaagcac	agctgtatat	aatggtgatg	tcataataaa	gttgttttat	cttttaagta	2580
aaaaaaaaa	aaaaaa					2596
<210> 100						
<211> 1020						
<212> DNA						
<213> Homo	sapiens					
<400> 100						
aaactagggg	aaaatgtagc	caacatatac	aaagatcttc	agaaactctc	tegeetettt	60
aaagaccagc	tggtgtatcc	tcttctggct	tttacccgac	aagcactgaa	cctaccagat	120
	tggtcgtcct					180
gttcgttccg	tcttgtcttt	gtctgcggtt	tgtcgtgacc	tctttactgc	ttcaaatgac	240
	ggaggtttt					300
	ggaagactgt					360
	atgctcctgc					420
	ccatttccta					480
	acacttccct					540
	agccagtttc					600
	cccatcttgc					660
	cggccaactg					720
	ccatttgttt					780
	-	-				

```
tctcgagtgt tattttctga ttgtggtgtt gagagttgca ctcccagaaa ccttttaaga
                                                                      840
gatacattta tagccctagg ggtggtatga cccaaaggtt cctctgtgac aaggttggcc
                                                                      900
ttgggaatag ttggctgcca atctccctgc tcttggttct cctctagatt gaagtttgtt
                                                                      960
                                                                     1020
<210> 101
<211> 1520
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (71)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (473)
<223> n equals a,t,g, or c
<400> 101
gcttttttct taagtgcaca aagcatcata ctccctggag gcaaacacat cgggctgctt
                                                                       60
cagcgttacg ngatgcttag cattttgaat attgtggcaa aaaaattaaa agttcactta
                                                                      120
ttaatattta tcagcagtat cataatttcc atcctctat ttcagaattt cacttgaggc
                                                                      180
aaaaatacca caagtgtaat tactctagca cagctattaa tgtgctggat gataggccac
                                                                      240
tgcgtcacat gaccttctat tgttcatggg tttaaagaga aagcagggct ttgtatttct
                                                                      300
ttttcttctt ttaaagtcga ctgtagcatc ttggcttttg tctggggtgg ggaggatctg
                                                                      360
                                                                      420
gggtctggtt cactttgtaa aagtaaacca tgtctgttta aacaatagag gtgtttaaga
agactettta gtttteetge agattgttea agattacatg ataateacae ggngtattta
                                                                      480
                                                                      540
tttcctactg acaaaccaag tacttgttac atcaccaatg gtaccaggag atgaagacgc
gggttttgag caggagcgag attaccaccc aaaaagggag ctacctgagg cagcccagct
                                                                      600
tctagcaaac tttttacatg ttgcacattt cagttcttaa atgaaggcta ctccagtgtc
                                                                      660
atttcattaa agtacctggg tgtagtactc aagtcccccc tcaagagttc ataagtaagc
                                                                      720
agtateettt tggecagtgg teetgttttt geecetaeee agaetgtteg rgaageatat
                                                                     780
                                                                      840
tctatagata aatctgacat ttgtcatcca ataccattgc agtcctctgc agcatacatt
ctcaatgggg gctgtatcac ccctagattg gttctgagat actgcaatgt cttgtgtcct
                                                                      900
tccaaaggac cataatactt gagcaaatat gaacatttct tggggtgagg gcagaaagag
                                                                     960
                                                                    1020
agaaacaaaa gtctaaaaag ggacaataat gaaaaaacag ttgagacctt tagtatgatg
                                                                    1080
ggaacaggat gaggaaggag gagatactga caggagccct gggtcttgct ctgcattaaa
cagatattta tggacattaa acagatattt atggagcacg actctgtacc ctacaggccc
                                                                    1140
agaatagtct taaggctcct gggaattgat gataggccat ttacccagtt tcagtttaga
                                                                    1200
ggcagattca ctggccttag catttcagta attatattta tttatttta gcctgaacca
                                                                    1260
gatttaatag gagaaactac tttctgcgtt tcttttaatt acttgtagtt tacacagtaa
                                                                    1320
ctttagaaga gtaaatgaaa gcatgcttcg atgctgccac tgtaaatacc attcattagt
                                                                    1380
aacttatttt ccctggagtc ttgtgaagtg tgaatttaaa gcctgctcta tctggaatat
                                                                    1440
ggaatagtat taagattaca agcacatttt atattcatga gccggaaagg caaaaaaaaa
                                                                    1500
aaaaaaaat gaccctcgag
                                                                    1520
<210> 102
```

<211> 1306

<212> DNA

<213> Homo sapiens

```
<220>
<221> SITE
<222> (1300)
<223> n equals a,t,g, or c
<400> 102
aattcccggg tcgacccacg cgtccggaat ttaagggacc cacactacct tcccgaagtt
                                                                       60
gaaggcaagc ggtgattgtt tgtagacggc gctttgtcat gggacctgtg cggttgggaa
                                                                      120
                                                                      180
tattgctttt ccttttttg gccgtgcacg aggcttgggc tgggatgttg aaggaggagg
                                                                      240
acgatgacac agaacgcttg cccagcaaat gcgaagtgtg taagctgctg agcacagagc
                                                                      300
tacaggegga actgagtege aceggtegat ctegagaggt getggagetg gggcaggtge
                                                                      360
tggatacagg caagaggaag agacacgtgc cttacagcgt ttcagagaca aggctggaag
aggeettaga gaatttatgt gageggatee tggaetatag tgtteaeget gagegeaagg
                                                                      420
gctcactgag atatgccaag ggtcagagtc agaccatggc aacactgaaa ggcctagtgc
                                                                      480
agaagggggt gaaggtggat ctggggatcc ctctggagct ttgggatgag cccagcgtgg
                                                                      540
aggtcacata cctcaagaag cagtgtgaga ccatgttgga rgargaggar gaagaggagg
                                                                      600
aagaggaagg gggagacaag atgaccaaga caggaagcca ccccaaactt gaccgagaag
                                                                      660
atctttgacc cttgcctttg agcccccagg aggggaaggg atcatggaga gccctctaaa
                                                                      720
gcctgcactc tccctgctcc acagctttca gggtgtgttt atgagtgact ccacccaagc
                                                                      780
ttgtagctgt tctctcccat ctaacctcag gcaagatcct ggtgaaacag catgacatgg
                                                                      840
cttctggggt ggaggtggg ggtggaggtc ctgctcctag agatgaactc tatccagccc
                                                                      900
cttaattggc aggtgtatgt gctgacagta ctgaaagctt tcctctttaa ctgatcccac
                                                                      960
ccccacccaa aagtcagcag tggcactgga gctgtgggct ttggqqaaqt cacttagctc
                                                                     1020
                                                                     1080
cttaaggtct gtttttagac ccttccaagg aagaggccag aacggacatt ctctgcgatc
                                                                     1140
tatatacatt gcctgtatcc aggaggctac acaccagcaa accgtgaagg agaatgggac
actgggtcat ggcctggagt tgctgataat ttaggtggga tagatacttg gtctacttaa
                                                                     1200
                                                                     1260
gctcaatgta acccagagcc caccatatag ttttataggt gctcaatttt ctatatcgct
1306
<210> 103
<211> 785
<212> DNA
<213> Homo sapiens
<400> 103
cttttagaag gtacgcctgc aggtaccggt ccggaattcc cgggtcgacc cacgcgtccg
                                                                      60
ggaaatgaac taccatttat aacttctgtt tttttattga gaaaatgatt cacgaattcc
                                                                      120
                                                                      180
aaatcagatt gccaggaaga aataggacgt gacggtactg ggccctgtga ttctcccagc
                                                                      240
ccttgcagtc cgctaggtga gaggaaaagc tctttacttc cgcccctggc agggacttct
                                                                      300
gggttatggg agaaaccaga gatgggaatg aggaaaatat gaactacagc agaagcccct
                                                                      360
gggcagctgt gatggagccc ctgacattac tcttcttgca tctgtcctgc cttctttccc
tctgcgaggc agtggggtgg gattcagagt gcttagtctg ctcactggga gaagaagagt
                                                                      420
tcctgcgcat gcaagccctg ctgtgtggct gtcgtttaca tttgggaggt gtcctgtatg
                                                                      480
tctgtacgtt ggggactgcc tgtatttgga agatttaaaa acctagcatc ctgttctcac
                                                                      540
cctctaagct gcattgagaa atgactcgtc tctgtatttg tattaagcct taacactttt
                                                                      600
                                                                      660
cttaagtgca ttcggtgcca acattttta gagctgtacc aaaacaaaaa gcctgtactc
acatcacaat gtcattttga taggagcgtt ttgttatttt tacaaggcag aatggggtgt
                                                                     720
aacagttgaa ttaaacttag caatcacgtg ctcaaaaaaa aaaaaaaaa aaaaagggcg
                                                                     780
gccgc
                                                                     785
```

<210> 104 <211> 2015

<212> DNA

60

```
<213> Homo sapiens
<220>
<221> SITE
<222> (3)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (9)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (1981)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (1990)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (2001)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (2002)
<223> n equals a,t,g, or c
<400> 104
conggaatno ogggtogaco cacgogtoog gootgogotg coagcagoca ggagocagga
                                                                         60
gccaagagca gagcgccagc atgaacttgg gggtcagcat gctgaggatc ctcttcctcc
                                                                        120
tggatgtagg aggagctcaa gtgctggcaa caggcaagac ccctggggct gaaattgatt
                                                                        180
tcaagtacgc cctcatcggg actgctgtgg gtgtcgccat atctgctggc ttcctggccc
                                                                        240
tgaagatctg catgatcagg aggcacttat ttgacgacga ctcttccgac ctgaaaagca
                                                                        300
creetggggg ceteagtgae accatecege taaagaagag ageeccaagg egaaaceaca
                                                                        360
atttctccaa aagagatgca caggtgattg agctgtaggt gagcagtgac gtgaagaggg
                                                                        420
gttctagccc cgtggaaaac agcccatggt taacatctca ggatgtcctg cattcaaaca
                                                                        480
cccaaggctg gtaatgaact ttcacatgga ctgaatattg gaggcaaata atagaaggaa
                                                                        540
                                                                        600
tagaatatac agtgcctctg tcctgaagga aaatatcatg cctcttctgg aagaaacgga
ctgcacagag gaaggattga gcaatttagc ctgcagtgga agaaggtgga caccaaaagc
                                                                        660
                                                                        720
ttcaccctgt gttggagctg ttcatgcttc catgaggcca tggtgtccat gtccgtggaa
cctaccacag aaaatggctc atgaaaaggg gaatccgacc caacacacag cttcctacac
                                                                        780
actgccatct tatcaacagt taggcactac tttgtagaac gattagcttc accctcttag
                                                                        840
ctgccaggag atcccttctt aaagatggac tatgtgaaga ttcgggagtc ctgaaacatg
                                                                        900
gggactccgg gatggtctct agccctatcg atgatgaaca ctggccttct ggaggggaaa
                                                                        960
tggcagtctg ggctggcgtg gtaggaaggg ctttggtgtt catggaatgg gcctgctgct
                                                                       1020
                                                                       1080
ctcagacctt caaaggatgg aaccaacgaa ggaccaaatg agaaagcaga tgcttgcctt
gcagagggcc atgaatgtca gttattattt ttctccttat acaattattt tgtggttatt
                                                                       1140
attacaatgt acatggctgt tgcatagaag acatgactgg tggaggctga ggaaagccat
                                                                       1200
gacattetac aattgccatc aggetaagge cccgtgagca ttteteccc ttgtaatatt
                                                                       1260
aaccctgtat ttctgggatc acatcacgga atattctttg cctttccact ttccaggaaa
                                                                       1320
```

and acres of

tctctcggac tgggctaccc tccttgtgtg tgatgaaaga tgagctatat ttcagaacaa agtgctgtgt tgtcatratt tgcctggact cccagggcgt ctcttaccca acttgataac gatgctgttc attagcagcc tttgttaact gataaccaag agcggtaatg tgatactcat aagcaatttt ctgtgtgtag gataaaataa accatcttgt atgggaaaaa aaaaaaaaaa	1380 1440 1500 1560 1620 1680 1740 1800 1960 1920 1980 2015
<210> 105 <211> 367 <212> DNA <213> Homo sapiens	
<pre><400> 105 cggcacgagt gtaaatgtca ccaccaaagg tttgcaccct gatcaaaaag agtatgaaaa gaataatacc acaacactta tggcctgtct tggaggcctt ctggggatta ttggtgtgat atgtcttatc agctgcctct ctccagaaat gaactgtgat ggtggacaca gctatgtgag gaattactta cagaaaccaa cctttgcatt aggtgágctt tatcctcctc tgataaatct ctgggaagca ggaaaagaaa aaagtacatc actgaaagta aaagcaactg ttataggttt accaacaaat atgtcctaaa aaccaccaag gaaacctact ccaaaaatga aaaaaaaaa aaaaaaaa</pre>	60 120 180 240 300 360 367
<210> 106 <211> 1889 <212> DNA <213> Homo sapiens	
<211> 1889 <212> DNA <213> Homo sapiens <400> 106 ctcatccttc tatcatcata tggagtggca ataatgaaaa tgaggaggcg ctgatgatga	60 120
<211> 1889 <212> DNA <213> Homo sapiens <400> 106 ctcatccttc tatcatcata tggagtggca ataatgaaaa tgaggaggcg ctgatgatga attggtatca tatcagtttc actgaccggc caatctacat caaggactat gtgacactct	
<211> 1889 <212> DNA <213> Homo sapiens <400> 106 ctcatccttc tatcatcata tggagtggca ataatgaaaa tgaggaggcg ctgatgatga	120
<211> 1889 <212> DNA <213> Homo sapiens <400> 106 ctcatccttc tatcatcata tggagtggca ataatgaaaa tgaggaggcg ctgatgatga attggtatca tatcagtttc actgaccggc caatctacat caaggactat gtgacactct atgtgaaaaa catcagagag ctcgtactgg caggagacaa gagtcgtcct tttattacgt	120 180
<211> 1889 <212> DNA <213> Homo sapiens <400> 106 ctcatccttc tatcatcata tggagtggca ataatgaaaa tgaggaggcg ctgatgatga attggtatca tatcagtttc actgaccggc caatctacat caaggactat gtgacactct atgtgaaaaa catcagagag ctcgtactgg caggagacaa gagtcgtcct tttattacgt ccagtcctac aaatggggct gaaactgttg cagaagcctg ggtctctcaa aaccctaata gcaattatt tggtgatgta cattttatg actatatcag tgattgctgg aactggaaag	120 180 240
<211> 1889 <212> DNA <213> Homo sapiens <400> 106 ctcatccttc tatcatcata tggagtggca ataatgaaaa tgaggaggcg ctgatgatga attggtatca tatcagttc actgaccggc caatctacat caaggactat gtgacactct atgtgaaaa catcagagag ctcgtactgg caggagacaa gagtcgtcct tttattacgt ccagtcctac aaatggggct gaaactgttg cagaagcctg ggtctctcaa aaccctaata gcaattatt tggtgatgta catttttatg actatatcag tgattgctgg aactggaaag ttttcccaaa agctcgattt gcatctgaat atggatatca gtcctggccg tccttcagta	120 180 240 300
<211> 1889 <212> DNA <213> Homo sapiens <400> 106 ctcatccttc tatcatcata tggagtggca ataatgaaaa tgaggaggcg ctgatgatga attggtatca tatcagtttc actgaccggc caatctacat caaggactat gtgacactct atgtgaaaaa catcagagag ctcgtactgg caggagacaa gagtcgtcct tttattacgt ccagtcctac aaatggggct gaaactgttg cagaagcctg ggtctctcaa aaccctaata gcaattatt tggtgatgta cattttatg actatatcag tgattgctgg aactggaaag	120 180 240 300 360
<211> 1889 <212> DNA <213> Homo sapiens <400> 106 ctcatccttc tatcatcata tggagtggca ataatgaaaa tgaggaggcg ctgatgatga attggtatca tatcagttc actgaccggc caatctacat caaggactat gtgacactct atgtgaaaa catcagagag ctcgtactgg caggagacaa gagtcgtcct tttattacgt ccagtcctac aaatggggct gaaactgttg cagaagcctg ggtctctcaa aaccctaata gcaattatt tggtgatgta catttttatg actatatcag tgattgctgg aactggaaag ttttcccaaa agctcgatt gcatctgaat atggatatca gtcctggccg tccttcagta cattagaaaa ggtctcgtct acagaggact ggtctttcaa tagcaagttt tcacttcatc	120 180 240 300 360 420
<211> 1889 <212> DNA <213> Homo sapiens <400> 106 ctcatccttc tatcatcata tggagtggca ataatgaaaa tgaggaggcg ctgatgatga attggtatca tatcagttc actgaccggc caatctacat caaggactat gtgacactct atgtgaaaa catcagagag ctcgtactgg caggagacaa gagtcgtcct tttattacgt ccagtcctac aaatggggct gaaactgttg cagaagcctg ggtctctaa aaccctaata gcaattatt tggtgatga catttttatg actatatcag tgattgctgg aactggaaag ttttcccaaa agctcgatt gcatctgaat atggatatca gtcctggccg tccttcagta cattagaaaa ggtctcgtct acagaggact ggtctttcaa tagcaagttt tcacttcatc gacaacatca cgaaggtgt aacaaacaaa tgctttatca ggctggactt catttcaaac	120 180 240 300 360 420 480 540
<pre><211> 1889 <212> DNA <213> Homo sapiens <400> 106 ctcatccttc tatcatcata tggagtggca ataatgaaaa tgaggaggcg ctgatgatga attggtatca tatcagttc actgaccggc caatctacat caaggactat gtgacactct atgtgaaaaa catcagagag ctcgtactgg caggagacaa gagtcgtcct tttattacgt ccagtcctac aaatggggct gaaactgttg cagaagcctg ggtctctaa aaccctaata gcaattattt tggtgatgta catttttatg actatatcag tgattgctgg aactggaaag ttttcccaaa agctcgatt gcatctgaat atggatatca gtcctggcg tccttcagta cattagaaaa ggtctcgtct acagaggact ggtctttcaa tagcaagttt tcacttcatc gacaacatca cgaaggtgt aacaaacaaa tgctttatca ggctggactt catttcaaac tcccccaaag cacagatcca ttacgcacat ttaaagatac catctacctt actcaggtga</pre>	120 180 240 300 360 420 480 540
<pre><211> 1889 <212> DNA <213> Homo sapiens <400> 106 ctcatcette tateateata tggagtggea ataatgaaaa tgaggaggeg ctgatgatga attggtatea tateagttte actgacegge caatetacat caaggaetat gtgacactet atgtgaaaaa cateagagag ctcgtactgg caggagacaa gagtegteet tttattacgt ecagteetae aaatgggget gaaactgttg cagaageetg ggtetetaa aaceetaata geaattatt tggtgatga catttttatg actatateag tgattgetgg aactggaaag tttteecaaa agetegatt geatetgaat atggatatea gteettggeeg teetteagta cattagaaaa ggtetegtet acagaggaet ggtetteaa tageaagtt teaetteate gacaacatea egaaggtgt aacaaacaaa tgetttatea ggetggaett cattteaaac teeeceaaag cacagatea ttaegeacat ttaaagatae catetacett acteaggtga tgeaggeeca gtgtgteaaa acagaactg aattetaceg cegtagtege agegagatag</pre>	120 180 240 300 360 420 480 540 600 660 720
<pre><211> 1889 <212> DNA <213> Homo sapiens <400> 106 ctcatcette tateateata tggagtggca ataatgaaaa tgaggaggeg ctgatgatga attggtatea tateagtte actgacegge caatetacat caaggactat gtgacactet atgtgaaaaa cateagagag ctcgtactgg caggagacaa gagtegteet tttattacgt ccagteetae aaatgggget gaaactgttg cagaageetg ggtetetaa aaccetaata gcaattatt tggtgatgta cattttatg actatateag tgattgetgg aactggaaag tttteccaaa agetegatt gcatetgaat atggatatea gteettggeeg teetteagta cattagaaaa ggtetegtet acagaggaet ggtetteaa tageaagtt teaetteate gacaacatea cgaaggtgt aacaaacaaa tgetttatea ggetggaett cattteaaac teeeccaaag cacagateca ttacgcacat ttaaagatae catetacett acteaggtga tgcaggeeca gtgtgteaaa acagaaactg aattetaceg cegtagtege agegagatag tggateagea agggecaag atgggggeac tttattggca gttgaatgac atetggcaag</pre>	120 180 240 300 360 420 480 540 600 720 780
<pre><211> 1889 <212> DNA <213> Homo sapiens <400> 106 ctcatcette tateateata tggagtggca ataatgaaaa tgaggaggeg ctgatgatga attggtatea tateagtte actgacegge caatetacat caaggactat gtgacactet atgtgaaaaa cateagagag ctcgtactgg caggagacaa gagtegteet tttattacgt ccagteetae aaatgggget gaaactgttg cagaageetg ggtetetaa aaccetaata gcaattatt tggtgatgta cattttatg actatateag tgattgetgg aactggaaag ttteecaaa agetegatt gcatetgaat atggatatea gteettggeagetg teetteagta cattagaaaa ggtetegtet acagaggaet ggtetteaa tageaagtt teaetteate gacaacatea cgaaggtgt aacaaacaaa tgetttatea ggetggaett cattteaaac teeeccaaag cacagateea ttacgcacat ttaaagatae catetacett acteaggtga tggateagea agggecea gtgtgteaaa acagaaactg aattetaceg cegtagtege agegagatag tggateagea agggecacag atgggggeae tttattggea gttgaatgae atettggeaag eteetteetg ggetteettt gatacggagg aaagtggaaa atgetteatt actttgetea</pre>	120 180 240 300 360 420 480 540 600 720 780 840
<pre><211> 1889 <212> DNA <213> Homo sapiens <400> 106 ctcatccttc tatcatcata tggagtggca ataatgaaaa tgaggaggcg ctgatgatga attggtatca tatcagtttc actgaccggc caatctacat aggagtcatc tattacgt caagtcctac aaatggggct gaaactgttg caggagacaa gggtcgtcct tttattacgt cagtcctac aaatggggct gaaactgttg cagaagcctg ggtctctaa aaccctaata gcaattattt tggtgatga cattttatg actatacag tggttgtcgaaag ttttcccaaa agctcgatt gcatctgaat atggatatca gtcctggcg tccttcagta cattagaaaa ggtccgtct acagaggact ggtctttaa atggatatca gtcctggcag tccttcagta cattagaaaa ggtctcgtct acagaggact ggtctttcaa tagcaagttt tcacttcatc gacaacatca cgaaggtggt aacaacaaaa tgctttatca ggctggactt cattcaacac tcccccaaag cacagatcca ttacgcacat ttaaagatac catctacctt actcaggtga tggatcagca agggcacacg atgggggacac tttattggca gttgaatgac atctggcaag ctccttcctg ggcttctctt ggatacggagg aaagtggaaa atgcttcatt acttggcaag gttgtgtcaga tcttcactcg gattattcga tggcagactc tttgaaatga ggaggagggggggggg</pre>	120 180 240 300 360 420 480 540 600 720 780 840 900
<pre><211> 1889 <212> DNA <213> Homo sapiens <400> 106 ctcatcettc tatcatcata tggagtggca ataatgaaaa tgaggaggcg ctgatgatga attggtatca tatcagtttc actgaccggc caatctacat caaggactat gtgacactct atgtgaaaaa catcagagag ctcgtactgg cagagacaa gagtcgtcet tttattacgt cagtcetac aaatggggct gaaactgttg cagaagcetg ggtctctaa aaccctaata gcaattattt tggtgatgta cattttatg actatatcag tgattgctgg aactggaaag tttcccaaa agctcgattt gcatctgaat atggatatca gtcctggcg tccttcaata accgaaag tttcccaaa agctcgattt gcatctgaat atggatatca gtcctggcg tccttcaata accgaaag tattagaaaa ggtctcgtct acagaggact ggtcttcaa tagcaagttt tcacttcatc gacaacatca cgaaggtggt aacaacaaaa tgctttatca ggctggactt cattcaacac tcccccaaag cacagatcca ttacgcacat ttaaagatac catctacctt actcaggtga tggatcagca agggcacacg atgggggaca tttattggca gttgaatgac atctggcaag ctccttcctg ggcttctctt gatacggagg aaattctaccg ccgtagtcgc agcgagatag tggatcaga tcttcactg gatattcga tggcaacacag atgggggaa aaggggaaa atgcttcatt actttgctca gatttcttt gctccactgt tgccagtagc tttgaaatga aaacatgttc tatatctatg gtgtgtcaga tcttcactcg gattattcga tggacactcag tgtgagagtc catacatgga gctccttgga gcccgtgtgc tctctgtgtga ctgaacgttt tgtgatgaaa ggaggagagg ctgtctcgct ttatgaggag ccagtgtctg aattgctaga gagatgtggg aattgcacac</pre>	120 180 240 300 360 420 480 540 600 720 780 840 900
<pre><211> DNA <213> Homo sapiens <400> 106 ctcatccttc tatcatcata tggagtggca ataatgaaaa tgaggaggcg ctgatgatga attggtatca tatcagtttc actgaccggc caatctacat caaggactat gtgacactct atgtgaaaaa catcagagag ctcgtactgg cagagacaa gagtcgtcct tttattacgt cagtcctac aaatggggct gaaactgttg cagaagcctg ggtctctaa aaccctaata gcaattatt tggtgatgta cattttatag actatatcag tgattgctgg aactggaaag tttcccaaa agctcgattt gcatctgaat atggatatac gtcattgcagaa ggtctgtct tcctaa aaccctaata gcaattattt tggtgatgta cattttatag actatatcag tgattgctgg aactggaaag tattaccaaaa agctcgattt gcatctgaat atggatatac gtcctggccg tccttcagta cattagaaaa ggtctcgtc acagagacct ggtctttcaa tagcaagttt tcacttcatc gacaacatca cgaaggtggt aacaaacaaa tgctttatca ggctggactt cattcaaac tcccccaaag cacagatcca ttacgcacat ttaaagatac catctacctt actcaggtga tggagagcacag atgggggacactg aattctaccg ccgtagtcc agcgagatag tggatcagca agggcacacg atgggggac ttattggca gttgaatgac atctggcaag ctccttcctg ggcttctctt gatacggagg aaagtggaaa atgcttcatt actttgctca gaatttcttt gctccactgt tgccagtagc tttgaaatga aacaatgtc tatatctatg gtgtgtcaga tcttcactcg gattattcga tggacactcag tgtgagagt catacatgga gctccttgac ttatgaggag ccagtgtctg aattgctgag gagatgtgg aattgcacac gggaaagctg tgtggtttcc ttttaccttt cagctgacca tgaacctct agacccac agggaaagctg tgtggtttcc ttttaccttt cagctgacca tgaacctctg agcccgacca tgaaccac gggaaagctg tgtggtttcc ttttaccttt cagctgacca tgaacctctg agcccgacca tgaaccac tgaacctct agccgaccaccac tgaaccaccac tgaaccaccaccaccaccaccaccaccaccaccaccaccac</pre>	120 180 240 300 360 420 480 540 600 720 780 840 900 960
<pre><211> 1889 <212> DNA <213> Homo sapiens <400> 106 ctcatccttc tatcatcata tggagtggca ataatgaaaa tgaggaggcg ctgatgatga attggtatca tatcagtttc actgaccggc caatctacat caaggactat gtgacactct atgtgaaaaa catcagagag ctcgtactgg caggagacaa gagtcgtcct tttattacgt ccagtcctac aaatggggct gaaactgttg cagaagcctg ggtctctcaa aaccctaata gcaattattt tggtgatgta cattttatg actatatcag tgattgctgg aactggaaag tttcccaaa agctcgatt gcatctgaat atggatatca gtccttgcag aactggaagg tcgtcacaa agctcgatt gcatctgaat atggatatca gtcctggcg tccttcaatc gacaacatca cgaaggtggt aacaacaaca tggttttcaa tagcaagttt tcaccttcatc gacaacatca cgaaggtggt aacaaacaaa tgctttatca ggctggactt catttcaact tcccccaaag cacagacca ttacgcacat ttaaagaatac catctacctt actaggtgg tggatcagaa acaggacca attattcacg tgtgatgac attatggcag tttattggca gttgaatgac atctggcaag ctccttcctg ggcttctctt gatacggaggagaaa atgcttcatt acttggcaag ctccttcctg ggcttctctt gatacggagg aaagtggaaa atgcttcatt acttgctca gaatttcttt gctccactgt tgccagtagc tttgaaatga aaccatgttc tatatctatg gtgtgtcaga tcttcactcg gattattcga tggacactcag tggagaggg ctgtctgcct ttataggagg ccagtgtcg actgaacgtt tgtgatgaaa ggaggagagg ctgtctgcct ttatagaggag ccagtgtctg aattgctaga gagatgggg aattgcacac gggaaagctg tgtggttcc ttttaccttt cagctgagc tcgaacgcca tgaacccca actacctg cagaccca ccacacct cctgtcctca ccgaaggagg ccgtggggc ccgtgggct cctgaaggcg cagatcactg cagaccca actaccctt cctgtcctc ccgaaggagg ccgtggggcc ccgtgggct cctgtagcca tgaacccac accacct cctgaaccac ccgaaggagg ccgtggggc cctgtcctcaccccacct cctgtcctc ccgaaggagg ccgtggggcc cctgtagcccaccaccaccaccaccaccaccaccaccaccaccac</pre>	120 180 240 300 360 420 480 540 600 720 780 840 900 960 1020 1080
<pre><211> DNA <213> Homo sapiens <400> 106 ctcatccttc tatcatcata tggagtggca ataatgaaaa tgaggaggcg ctgatgatga attggtatca tatcagtttc actgaccggc caatctacat caaggactat gtgacactct atgtgaaaaa catcagagag ctcgtactgg cagagacaa gagtcgtcct tttattacgt cagtcctac aaatggggct gaaactgttg cagaagcctg ggtctctaa aaccctaata gcaattatt tggtgatgta cattttatag actatatcag tgattgctgg aactggaaag tttcccaaa agctcgattt gcatctgaat atggatatac gtcattgcagaa ggtctgtct tcctaa aaccctaata gcaattattt tggtgatgta cattttatag actatatcag tgattgctgg aactggaaag tattaccaaaa agctcgattt gcatctgaat atggatatac gtcctggccg tccttcagta cattagaaaa ggtctcgtc acagagacct ggtctttcaa tagcaagttt tcacttcatc gacaacatca cgaaggtggt aacaaacaaa tgctttatca ggctggactt cattcaaac tcccccaaag cacagatcca ttacgcacat ttaaagatac catctacctt actcaggtga tggagagcacag atgggggacactg aattctaccg ccgtagtcc agcgagatag tggatcagca agggcacacg atgggggac ttattggca gttgaatgac atctggcaag ctccttcctg ggcttctctt gatacggagg aaagtggaaa atgcttcatt actttgctca gaatttcttt gctccactgt tgccagtagc tttgaaatga aacaatgtc tatatctatg gtgtgtcaga tcttcactcg gattattcga tggacactcag tgtgagagt catacatgga gctccttgac ttatgaggag ccagtgtctg aattgctgag gagatgtgg aattgcacac gggaaagctg tgtggtttcc ttttaccttt cagctgacca tgaacctct agacccac agggaaagctg tgtggtttcc ttttaccttt cagctgacca tgaacctctg agcccgacca tgaaccac gggaaagctg tgtggtttcc ttttaccttt cagctgacca tgaacctctg agcccgacca tgaaccac tgaacctct agccgaccaccac tgaaccaccac tgaaccaccaccaccaccaccaccaccaccaccaccaccac</pre>	120 180 240 300 360 420 480 540 600 720 780 840 900 960

PCT/US99/05804

WO 99/47540

62

```
tgactgagaa gacacgaact atattatttt accettggga geecaccage aagaatgagt
                                                                      1260
tggagcaatc ttttcatgtg acctccttaa cagatattta ctgaaggaat ctaggttgta
                                                                      1320
                                                                      1380
ttttcagtgg acaatgggaa taaagcattt ctaaagcacc gactggagag gaaggcaaca
gagacaagga gagaagccga gagacatgtc tgcgtgctgc cacgcatctg agcgattgct
                                                                      1440
ctgtgaagag ttgtacactg aacattttca ggggaggctg tttacccagg caatgtcctc
                                                                      1500
aaacaagcct gtgccggggt gtcctggaat ctgtgccagg actgtgtttt tagcccttca
                                                                      1560
                                                                      1620
                                                                      1680
                                                                      1740
                                                                      1800
                                                                      1860
                                                                      1889
```

```
cctctcagct ttagcaggac atgaaccagt tataacaaga tggccctgca gctggttaca
 agaatgtgac atggcaggat ctatggaacc aaatggaagg ttttgaggtg atgtaggtct
 ttcacagtta gctttgggga atacagaata ctcaaataaa gtgctttgtt attatttcag
 agggaatggc gattgaaatg ttacaacaga gatttcttgg tggtagctat ttgggtaaag
 gtatatggat atttttctgt acatgtgaaa ttatataaaa ataaaagtta tataaattac
 attgaaaaaa aaaaaaaaa aaaaaaaaa
<210> 107
<211> 1201
<212> DNA
<213> Homo sapiens
<220>
<221> SITE
<222> (1086)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (1161)
<223> n equals a,t,g, or c
<220>
<221> SITE
<222> (1176)
<223> n equals a,t,g, or c
<400> 107
cggcacgagc ggctggcagc acgactcgcg taccgtgcgc cgattgcctc tcggcctggg
                                                                         60
                                                                        120
caatggtccc ggctgccggt cgacgaccgc cccgcgtcat gcggctcctc ggctggtggc
                                                                        180
aagtattgct gtgggtgctg ggacttcccg tccgcggcgt ggagggacct tatggatttt
                                                                        240
ctgaacccaa acggtagtga ctgtactcta gtcctgtttt acaccccgtg gtgccgcttt
                                                                        300
totgecagtt tggcccctca ctttaactct ctgccccggg catttccagc tcttcacttt
                                                                        360
ttggcactgg atgcatctca gcacagcagc ctttctacca ggtttggcac cgtagctgtt
                                                                        420
cctaatattt tattatttca aggagctaaa ccaatggcca gatttaatca tacagatcga
                                                                        480
acactggaaa cactgaaaat cttcattttt aatcagacag gtatagaagc caagaagaat
                                                                        540
gtggtggtaa ctcaagccga ccaaataggc cctcttccca gcactttgat aaaaagtgtg
                                                                        600
gactggttgc ttgtattttc cttattcttt ttaattagtt ttattatgta tgctaccatt
cgaactgaga gtattcggtg gctaattcca ggacaagagc aggaacatgt ggagtagtga
                                                                        660
tggtctgaaa gaagttggaa agaggaactt caatccttcg tttcagaaat tagtgctaca
                                                                        720
                                                                        780
gtttcataca ttttctccag tgacgtgttg acttgaaact tcaggcagat taaaagaatc
atttgttgaa caactgaatg tataaaaaaa ttataaactg gtgttttaac tagtattgca
                                                                        840
ataagcaaat gcaaaaatat tcaatagatg cactattctt gtttttactg catgmacgta
                                                                        900
atccagtatt tggkaaagta atccaktttg aaatgtgrag rtgtattccg gcagaatagt
                                                                        960
gagtagaatg acagettact atacagaagg cmaaaatagg actetcaggt aatagtttaa
                                                                       1020
ggaaaccctt gattccttat gatgatgttt aagaaaggtt agttttctgt ttctttgcca
gttttncttc taggagtcca tagccaggga aagtatgtga accagaattg gttagtgtga
cccctccaa gtagccagtg ntgggaaata agggtncaat accttgatgt ttgtgatctc
```

```
1201
 t
<210> 108
<211> 75
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (75)
<223> Xaa equals stop translation
<400> 108
Met Asp Pro Leu Cys Leu Pro Ile Ile Leu Phe Ser Ala Val Val Leu
Arg Asn Leu Phe His Leu Leu Ile Leu Thr Phe His Tyr Leu Pro Leu
Phe Cys Asp Asn Pro Leu Ile Leu Glu Asp Leu Ser Cys Ile His Leu
                            40
                                               45
Arg Val Asn Ile Phe Lys Ala Lys Gln Pro Lys Phe Tyr Gly Asn Gln
Leu Gln Pro Cys Val Met Lys Ser Ser Ala Xaa
                    70
<210> 109
<211> 202
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (202)
<223> Xaa equals stop translation
<400> 109
Met Lys Leu Leu Ile Leu Phe Leu Ser His Leu Leu Ser Leu Ala Phe
                                   10
Gly Ile Leu Cys Leu Ser Val Thr Val Ile Leu Ser Leu Leu Ser
                   25
Phe Ser Lys Arg Gly Phe Ser Val Arg Ser Phe Gly Thr Gly Thr His
Val Lys Leu Pro Gly Pro Ala Pro Asp Lys Pro Asn Val Tyr Asp Phe
                       55
Lys Thr Thr Tyr Asp Gln Met Tyr Asn Asp Leu Leu Arg Lys Asp Lys
                                        75
                   70
```

Glu Leu Tyr Thr Gln Asn Gly Ile Leu His Met Leu Asp Arg Asn Lys 95

Arg Ile Lys Pro Arg Pro Glu Arg Phe Gln Asn Cys Lys Asp Leu Phe 100

Asp Leu Ile Leu Thr Cys Glu Glu Arg Val Tyr Asp Gln Val Val Glu 115

Asp Leu Asn Ser Arg Glu Gln Glu Thr Cys Gln Pro Val His Val Val 130

Asn Val Asp Ile Gln Asp Asn His Glu Glu Glu Ala Thr Leu Gly Ala Phe 145

Leu Ile Cys Glu Leu Cys Gln Cys Ile Gln His Thr Glu Asp Met Glu

Asn Glu Ile Asp Glu Leu Leu Gln Glu Phe Glu Glu Lys Ser Gly Arg 180 185 190

170

Thr Phe Leu His Thr Val Cys Phe Tyr Xaa 195 200

<210> 110

<211> 371

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (31)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (193)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 110

Met Gly Leu Lys Leu Leu Gln Lys Pro Gly Ser Leu Lys Thr Leu Ile
1 5 10 15

Ala Ile Ile Leu Val Met Tyr Ile Phe Met Thr Ile Ser Val Xaa Cys 20 25 30

Trp Asn Trp Lys Val Phe Pro Lys Ala Arg Phe Ala Ser Glu Tyr Gly
35 40 45

Tyr Gln Ser Trp Pro Ser Phe Ser Thr Leu Glu Lys Val Ser Ser Thr

Glu Asp Trp Ser Phe Asn Ser Lys Phe Ser Leu His Arg Gln His His

65					70					75					80
Glu	Gly	Gly	Asn	Lys 85	Gln	Met	Leu	Tyr	Gln 90	Ala	Gly	Leu	His	Phe 95	Lys
Leu	Pro	Gln	Ser 100	Thr	Asp	Pro	Leu	Arg 105	Thr	Phe	Lys	Asp	Thr 110	Ile	Tyr
Leu	Thr	Gln 115	Val	Met	Gln	Ala	Gln 120	Cys	Val	Lys	Thr	Glu 125	Thr	Glu	Phe
Tyr	Arg 130	Arg	Ser	Arg	Ser	Glu 135	Ile	Val	Asp	Gln	Gln 140	Gly	His	Thr	Met
Gly 145	Ala	Leu	Tyr	Trp	Gln 150	Leu	Asn	Asp	Ile	Trp 155	Gln	Ala	Pro	Ser	Trp 160
Ala	Ser	Leu	Glu	Tyr 165	Gly	Gly	Lys	Trp	Lys 170	Met	Leu	His	Tyr	Phe 175	Ala
Gln	Asn	Phe	Phe 180	Ala	Pro	Leu	Leu	Pro 185	Val	Gly	Phe	Glu	Asn 190	Glu	Asn
Xaa	Phe	Tyr 195	Ile	Tyr	Gly	Val	Ser 200	Asp	Leu	His	Ser	Asp 205	Tyr	Ser	Met
Thr	Leu 210	Ser	Val	Arg	Val	His 215	Thr	Trp	Ser	Ser	Leu 220	Glu	Pro	Val	Cys
Ser 225	Arg	Val	Thr	Glu	Arg 230	Phe	Val	Met	Lys	Gly 235	Gly	Glu	Ala	Val	Cys 240
Leu	Tyr	Glu	Glu	Pro 245	Val	Ser	Glu	Leu	Leu 250	Arg	Arg	Cys	Gly	Asn 255	Cys
Thr	Arg	Glu	Ser 260	Cys	Val	Val	Ser	Phe 265	Tyr	Leu	Ser	Ala	Asp 270	His	Glu
Leu	Leu	Ser 275	Pro	Thr	Asn	Tyr	His 280	Phe	Leu	Ser	Ser	Pro 285	Lys	Glu	Ala
Val	Gly 290	Leu	Cys	Lys	Ala	Gln 295	Ile	Thr	Ala	Ile	Ile 300	Ser	Gln	Gln	Gly
Asp 305	Ile	Phe	Val	Phe	Asp 310	Leu	Glu	Thr	Ser	Ala 315	Val	Ala	Pro	Phe	Val 320
Trp	Leu	Asp	Val	Gly 325	Ser	Ile	Pro	Gly	Arg 330	Phe	Ser	Asp	Asn	Gly 335	Phe
Leu	Met	Thr	Glu 340	Lys	Thr	Arg	Thr	Ile 345	Leu	Phe	Tyr	Pro	Trp 350	Glu	Pro
Thr	Ser	Lys 355	Asn	Glu	Leu	Glu	Gln 360	Ser	Phe	His	Val	Thr 365	Ser	Leu	Thr

Asp Ile Tyr 370

<210> 111

<211> 114

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (38)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (114)

<223> Xaa equals stop translation

<400> 111

Met Arg Pro Leu Leu Gly Gly Tyr Trp Val Leu Cys Leu Ser Val

Leu Gly His Ala Ala Leu Tyr His Phe Trp Leu Arg Glu Glu Gly Lys 25

Gly Pro Pro Gln Val Xaa Ser Val Leu Ala Leu Ala Leu Pro Ala Gly 40

Ser Cys Ala Pro Gly Leu Pro Phe Pro Gly Pro Leu Ile Pro Thr Gln 55

Leu Leu Phe Ala Leu Glu Trp Gly Thr Pro Thr Pro Leu Arg Asp His 70

Pro Pro His Ser Met His Ser Ala Pro Gln Asn Pro Pro Val Phe Leu

Gly Thr His Thr Cys Pro Pro Ser Trp Tyr Phe Arg Leu Ile Pro Gln 105

Ala Xaa

<210> 112

<211> 152

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (152)

<223> Xaa equals stop translation

<400> 112

Met Arg Arg Leu Leu Val Thr Ser Leu Val Val Val Leu Leu Trp

Glu Ala Gly Ala Val Pro Ala Pro Lys Val Pro Ile Lys Met Gln Val

Lys His Trp Pro Ser Glu Gln Asp Pro Glu Lys Ala Trp Gly Ala Arg 35 45

Val Val Glu Pro Pro Glu Lys Asp Asp Gln Leu Val Val Leu Phe Pro

Val Gln Lys Pro Lys Leu Leu Thr Thr Glu Glu Lys Pro Arg Gly Gln 70 75

Gly Arg Gly Pro Ile Leu Pro Gly Thr Lys Ala Trp Met Glu Thr Glu

Asp Thr Leu Gly Arg Val Leu Ser Pro Glu Pro Asp His Asp Ser Leu 100 105 110

Tyr His Pro Pro Pro Glu Glu Asp Gln Gly Glu Glu Arg Pro Arg Leu 120

Trp Val Met Pro Asn His Gln Val Leu Leu Gly Pro Glu Glu Asp Gln 135 130 140

Asp His Ile Tyr His Pro Gln Xaa 145 150

<210> 113

<211> 56

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (56)

<223> Xaa equals stop translation

<400> 113

Met Pro Cys Gly Lys Phe Leu Phe Pro Val Ser Pro Val Ser Ser Leu

Ser Leu His Trp Ser Ala Val Leu Leu Leu Leu Ala Asp Phe Pro

Arg Val His Gly Ser Pro Pro Gly Val Ser Arg Val Ser Ile Leu His

Cys Leu Phe Pro Phe Leu Ser Xaa

68

<210> 114

<211> 237

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (237)

<223> Xaa equals stop translation

<400> 114

Met Glu Val Arg Leu Ile Phe Leu Ser Gly Leu Cys Ile Ala Val Ala 1 5 10 15

Val Val Trp Ala Val Phe Arg Asn Glu Asp Arg Trp Ala Trp Ile Leu 20 25 30

Gln Asp Ile Leu Gly Ile Ala Phe Cys Leu Asn Leu Ile Lys Thr Leu $35 \hspace{1.5cm} 40 \hspace{1.5cm} 45$

Lys Leu Pro Asn Phe Lys Ser Cys Val Ile Leu Leu Gly Leu Leu Leu 50 55 60

Leu Tyr Asp Val Phe Phe Val Phe Ile Thr Pro Phe Ile Thr Lys Asn 65 70 75 80

Gly Glu Ser Ile Met Val Glu Leu Ala Ala Gly Pro Phe Gly Asn Asn 85 90 95

Glu Lys Leu Pro Val Val Ile Arg Val Pro Lys Leu Ile Tyr Phe Ser 100 105 110

Val Met Ser Val Cys Leu Met Pro Val Ser Ile Leu Gly Phe Gly Asp 115 120 125

Ile Ile Val Pro Gly Leu Leu Ile Ala Tyr Cys Arg Arg Phe Asp Val 130 135 140

Gln Thr Gly Ser Ser Tyr Ile Tyr Tyr Val Ser Ser Thr Val Ala Tyr 145 150 155 160

Ala Ile Gly Met Ile Leu Thr Phe Val Val Leu Val Leu Met Lys Lys 165 170 175

Gly Gln Pro Ala Leu Leu Tyr Leu Val Pro Cys Thr Leu Ile Thr Ala 180 185 190

Ser Val Val Ala Trp Arg Arg Lys Glu Met Lys Lys Phe Trp Lys Gly 195 200 205

Asn Ser Tyr Gln Met Met Asp His Leu Asp Cys Ala Thr Asn Glu Glu 210 215 220

```
Asn Pro Val Ile Ser Gly Glu Gln Ile Val Gln Gln Xaa
                    230
<210> 115
<211> 44
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (44)
<223> Xaa equals stop translation
<400> 115
Met Phe Cys Phe Tyr Leu His Phe Ile Phe His Val Leu Ser Tyr Lys
                                     10
Leu Asn Pro Leu Leu Phe Phe Ser Cys Ser Cys Phe Cys Phe Ile Leu
            20
                                 25
Val Phe Leu Phe Pro Asp Tyr His Leu Gly Met Xaa
         35
<210> 116
<211> 65
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (65)
<223> Xaa equals stop translation
<400> 116
Met Val Arg His Ile Arg Glu Arg Arg Gln Pro Leu Ala Phe Gln
Arg Val Leu Leu Ser Leu Cys Leu Leu Glu Gly Ile Trp His Ser Pro
Ala Ala Ala Gly Gly Gly Ser His Cys Ser Ser Trp Pro Ser Leu
Tyr Thr Thr Phe Gln Arg Val Ser Leu Leu Glu Leu Asp Leu Gly Leu
                        55
Xaa
65
<210> 117
<211> 118
<212> PRT
```

70

```
<213> Homo sapiens
<220>
<221> SITE
<222> (118)
<223> Xaa equals stop translation
Met Ala Arg Ser Ala Leu Arg Leu Glu Ile Leu Gly Gln Leu Leu Val
                                     10
Gly Val Ser Ser Cys Cys Ala Glu Ile Arg Ser Arg Ser Tyr Leu Gly
Phe Cys Trp Lys Asn Ile Gln Asp Glu Arg Lys Lys Ile Ile Leu
Arg Gly Ser Arg Asn Leu Leu Cys Pro Arg Leu Leu Arg Pro Leu Glu
Pro Val Gln Ala Lys Gly Thr Gln Ser Val Asp Pro Arg Glu Val Val
Arg Glu Thr Arg Ser Met Ser Thr Leu Pro Ala Asp Phe Cys Leu Leu
Pro Gln Ala Ser Arg Met Ala Gln Lys Gly Ser Pro Ser Arg Ser Ser
Leu Gln Leu Leu Phe Xaa
       115
<210> 118
<211> 65
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (65)
<223> Xaa equals stop translation
<400> 118
Met Thr Val Ser Leu Phe Leu Leu Leu Ala Thr Ser Gln Ser Gln Asp
Gly Cys Cys Asp Ser Gly Ser Cys Pro Asn Ser Arg Gln Gln Glu Gly
His Gly Ala Ala Pro Ala Ser Arg Cys Pro Cys Arg Pro Ser Leu Gln
        35
                             40
                                                 45
```

Ala Gln Glu Pro Lys Glu Glu Ser Thr Gln Met Trp Cys Ser His Leu

60

55

71

Xaa 65

65

<210> 119

<211> 43

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (43)

<223> Xaa equals stop translation

<400> 119

Met Leu Lys Trp Thr Gly Phe Leu Val Val Leu Val Ala Phe Lys Lys

1 10 15

Ile Ser Ala Ser Phe Gln Val Asn Tyr Asn Leu Lys Phe Glu Ile Ser 20 25 30

Phe Gly Glu Pro Trp Lys Phe Thr Gln Trp Xaa 35

<210> 120

<211> 48

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (48)

<223> Xaa equals stop translation

<400> 120

Met Ser Phe Gly Ile Ser Ile His Thr Cys Thr Tyr Leu Ile Phe Ile
1 10 15

Ala Phe His Phe Ile Ala Leu Cys Lys Val Thr Phe Phe Thr Asp Ser 20 25 30

Arg Phe Gly Asn Pro Met Ser Ile Ser Leu Ser Ala Pro Phe Phe Xaa 35 40

<210> 121

<211> 140

<212> PRT

<213> Homo sapiens

```
<220>
<221> SITE
<222> (140)
<223> Xaa equals stop translation
Met Ala Leu Gly Ile Gln Lys Arg Phe Ser Pro Glu Val Leu Gly Leu
                                     10
Cys Ala Ser Thr Ala Leu Val Trp Val Val Met Glu Val Leu Ala Leu
Leu Leu Gly Leu Tyr Leu Ala Thr Val Arg Ser Asp Leu Ser Thr Phe
                             40
His Leu Leu Ala Tyr Ser Gly Tyr Lys Tyr Val Gly Met Ile Leu Ser
                        55
Val Leu Thr Gly Leu Leu Phe Gly Ser Asp Gly Tyr Tyr Val Ala Leu
Ala Trp Thr Ser Ser Ala Leu Met Tyr Phe Ile Val Arg Ser Leu Arg
                                     90
Thr Ala Ala Leu Gly Pro Asp Ser Met Gly Gly Pro Val Pro Arg Gln
            100
                                105
Arg Leu Gln Leu Tyr Leu Thr Leu Gly Ala Ala Phe Gln Pro Leu
                            120
Ile Ile Tyr Trp Leu Thr Phe His Leu Val Arg Xaa
<210> 122
<211> 92
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (89)
<223> Xaa equals any of the naturally occurring L-amino acids
<220>
<221> SITE
<222> (92)
<223> Xaa equals stop translation
Met Met Asp Phe Leu Arg Cys Val Thr Ala Ala Leu Ile Tyr Phe Ala
Ile Ser Ile Thr Ala Ile Ala Lys Tyr Ser Asp Gly Ala Ser Lys Ala
```

25

Ala Gly Gly Ser Val Pro Asp Thr Arg Ala Val Cys Pro Ser Arg Ser 35 40 45

Glu Met Gly Arg Glu Leu Gly Ala Ala Ser Arg Glu Gln Gly Val
50 55 60

Ser Pro Val Met His Pro Ile His Pro Val His Arg Cys Leu Ala Ser 65 70 75 80

Leu Leu Pro Ser Cys Leu Gln Leu Xaa Ser Thr Xaa 85 90

<210> 123

<211> 347

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (242)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (246)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (347)

<223> Xaa equals stop translation

<400> 123

Met Arg Arg Gly Ala Gly Ala Ala Arg Gly Arg Ala Ser Trp Cys Trp

1 10 15

Ala Leu Ala Leu Leu Trp Leu Ala Val Val Pro Gly Trp Ser Arg Val
20 25 30

Ser Gly Ile Pro Ser Arg Arg His Trp Pro Val Pro Tyr Lys Arg Phe

Asp Phe Arg Pro Lys Pro Asp Pro Tyr Cys Gln Ala Lys Tyr Thr Phe 50 55 60

Cys Pro Thr Gly Ser Pro Ile Pro Val Met Glu Gly Asp Asp Ile
65 70 75 80

Glu Val Phe Arg Leu Gln Ala Pro Val Trp Glu Phe Lys Tyr Gly Asp 85 90 95

Leu Leu Gly His Leu Lys Ile Met His Asp Ala Ile Gly Phe Arg Ser 100 105 110

Thr Leu Thr Gly Lys Asn Tyr Thr Met Glu Trp Tyr Glu Leu Phe Gln

Leu Gly Asn Cys Thr Phe Pro His Leu Arg Pro Glu Met Asp Ala Pro 135

Phe Trp Cys Asn Gln Gly Ala Ala Cys Phe Phe Glu Gly Ile Asp Asp 150 155

Val His Trp Lys Glu Asn Gly Thr Leu Val Gln Val Ala Thr Ile Ser 170 165

Gly Asn Met Phe Asn Gln Met Ala Lys Trp Val Lys Gln Asp Asn Glu 180 185

Thr Gly Ile Tyr Tyr Glu Thr Trp Asn Val Lys Ala Ser Pro Glu Lys 200

Gly Ala Glu Thr Trp Phe Asp Ser Tyr Asp Cys Ser Lys Phe Val Leu 215 220

Arg Thr Phe Asn Lys Leu Ala Glu Phe Gly Ala Glu Phe Lys Asn Ile 230

Glu Xaa Asn Tyr Thr Xaa Ile Phe Leu Tyr Ser Gly Glu Pro Thr Tyr 245 250

Leu Gly Asn Glu Thr Ser Val Phe Gly Pro Thr Gly Asn Lys Thr Leu 265

Gly Leu Ala Ile Lys Arg Phe Tyr Tyr Pro Phe Lys Pro His Leu Pro 280

Thr Lys Glu Phe Leu Leu Ser Leu Leu Gln Ile Phe Asp Ala Val Ile 295

Val His Lys Gln Phe Tyr Leu Phe Tyr Asn Phe Glu Tyr Trp Phe Leu 310 315

Pro Met Lys Phe Pro Phe Ile Lys Ile Thr Tyr Glu Glu Ile Pro Leu 330

Pro Ile Arg Asn Lys Thr Leu Ser Gly Leu Xaa 340 345

<210> 124

<211> 234

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (173)

75

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (234)

<223> Xaa equals stop translation

<400> 124

Met His Arg Gly Lys Leu Asp Cys Ala Gly Gly Ala Leu Leu Ser Ser 1 5 10 15

Tyr Leu Ile Val Leu Met Ile Leu Leu Ala Val Val Ile Cys Thr Val 20 25 30

Ser Ala Ile Met Cys Val Ser Met Arg Gly Thr Ile Cys Asn Pro Gly 35 40

Pro Arg Lys Ser Met Ser Lys Leu Leu Tyr Ile Arg Leu Ala Leu Phe 50 60

Phe Pro Glu Met Val Trp Ala Ser Leu Gly Ala Ala Trp Val Ala Asp
65 70 75 80

Gly Val Gln Cys Asp Arg Thr Val Val Asn Gly Ile Ile Ala Thr Val $85 \\ 90 \\ 95$

Val Val Ser Trp Ile Ile Ile Ala Ala Thr Val Val Ser Ile Ile Ile 100 105 110

Val Phe Asp Pro Leu Gly Gly Lys Met Ala Pro Tyr Ser Ser Ala Gly 115 120 125

Pro Ser His Leu Asp Ser His Asp Ser Ser Gln Leu Leu Asn Gly Leu 130 135 140

Lys Thr Ala Ala Thr Ser Val Trp Glu Thr Arg Ile Lys Leu Leu Cys 145 150 155 160

Cys Cys Ile Gly Lys Asp Asp His Thr Arg Val Ala Xaa Ser Ser Thr 165 170 175

Ala Glu Leu Phe Ser Thr Tyr Phe Ser Asp Thr Asp Leu Val Pro Ser 180 185 190

Asp Ile Ala Ala Gly Leu Ala Leu Leu His Gln Gln Gln Asp Asn Ile 195 200 205

Arg Asn Asn Gln Asp Leu Pro Arg Trp Ser Ala Met Pro Gln Gly Ala 210 215 220

Pro Arg Lys Leu Ile Trp Met Gln Asn Xaa 225 230

<210> 125

76

<211> 54 <212> PRT <213> Homo sapiens <220> <221> SITE <222> (54) <223> Xaa equals stop translation <400> 125 Met Gln Gly Val Leu Phe Gly Phe Val Trp Leu Phe Ser Phe Leu Trp 10 Gln Glu Asn Lys Ser Ser Ala Ser Pro Ser Thr Leu Ala Lys Ser Gly Ser Pro Cys Pro Val Ser Ile Pro Trp Met Pro Gly Val Leu Val Arg 40 Phe Phe Thr Leu Leu Xaa 50 <210> 126 <211> 82 <212> PRT <213> Homo sapiens <220> <221> SITE <222> (44) <223> Xaa equals any of the naturally occurring L-amino acids <220> <221> SITE <222> (82) <223> Xaa equals stop translation <400> 126 Met Arg Met Arg Val Ala Val Ala Pro Arg Pro His Gln His Leu Val Val Ser Val Ser Trp Ile Leu Ala Ile Leu Ile Ser Val Ser Gly Tyr 20 His Cys Phe His Leu Gln Phe Ser Tyr Met Val Xaa Asn Ile Phe Pro His Val Tyr Leu Ser Ser Ala Tyr Leu Leu Arg Pro Val Ile Cys Ser 50 55 Asp Leu Leu Pro Val Phe Val Cys Leu His Val Cys Leu Cys Leu Ile 65 70 75

Phe Xaa

PCT/US99/05804

```
<210> 127
<211> 42
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (42)
<223> Xaa equals stop translation
<400> 127
Met Gly Trp Glu Ala Ala Leu Ala Leu Leu Val Ser Ala Val Phe Phe
                  5
Pro Trp Cys Thr Ile Gln Arg Pro Asp Val Gly Thr Thr Ser Pro Gly
                                 25
Gly Leu Glu Arg Arg Ser Lys Gly Phe Xaa
<210> 128
<211> 66
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (66)
<223> Xaa equals stop translation
<400> 128
Met Thr Phe Met Ile Leu Lys Phe Phe Phe Leu Cys Gly Phe Val Leu
Asn Arg Leu Ile Ala Arg Gln Leu Ala Lys Ile His Ala Ile His Ala
His Asn Gly Trp Ile Pro Lys Ser Asn Leu Trp Leu Lys Met Gly Lys
Tyr Phe Ser Leu Ile Pro Thr Gly Phe Ala Asp Glu Asp Ile Asn Lys
Arg Xaa
 65
<210> 129
<211> 50
<212> PRT
<213> Homo sapiens
```

```
<220>
<221> SITE
<222> (50)
<223> Xaa equals stop translation
<400> 129
Met Ile Val Asn His Phe Ser Phe Leu Phe Cys Trp Ile Val Phe Cys
Phe Leu Leu Gln His Ser Cys Phe Cys Cys Ala Tyr Phe Trp Ser Phe
Asp Ser Leu Cys His Cys Phe Leu Ser His Thr Pro Leu Arg Phe Thr
                             40
Gln Xaa
     50
<210> 130
<211> 227
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (227)
<223> Xaa equals stop translation
<400> 130
Met Glu Thr Val Val Ile Val Ala Ile Gly Val Leu Ala Thr Ile Phe
Leu Ala Ser Phe Ala Ala Leu Val Leu Val Cys Arg Gln Arg Tyr Cys
Arg Pro Arg Asp Leu Leu Gln Arg Tyr Asp Ser Lys Pro Ile Val Asp
Leu Ile Gly Ala Met Glu Thr Gln Ser Glu Pro Ser Glu Leu Glu Leu
                         55
Asp Asp Val Val Ile Thr Asn Pro His Ile Glu Ala Ile Leu Glu Asn
Glu Asp Trp Ile Glu Asp Ala Ser Gly Leu Met Ser His Cys Ile Ala
                 85
                                    90
Ile Leu Lys Ile Cys His Thr Leu Thr Glu Lys Leu Val Ala Met Thr
            100
                                105
Met Gly Ser Gly Ala Lys Met Lys Thr Ser Ala Ser Val Ser Asp Ile
                            120
```

79

Ile Val Val Ala Lys Arg Ile Ser Pro Arg Val Asp Asp Val Val Lys
130 135 140

Ser Met Tyr Pro Pro Leu Asp Pro Lys Leu Leu Asp Ala Arg Thr Thr 145 150 155 160

Ala Leu Leu Ser Val Ser His Leu Val Leu Val Thr Arg Asn Ala 165 170 175

Cys His Leu Thr Gly Gly Leu Asp Trp Ile Asp Gln Ser Leu Ser Ala 180 185 190

Ala Glu Glu His Leu Glu Val Leu Arg Glu Ala Ala Leu Ala Ser Glu 195 200 205

Pro Asp Lys Gly Leu Pro Gly Pro Glu Gly Phe Leu Gln Glu Gln Ser 210 215 220

Ala Ile Xaa 225

<210> 131

<211> 118

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (118)

<223> Xaa equals stop translation

<400> 131

Met Gln Arg Ile Ala Ser Leu Leu Thr Leu Leu Thr Gln Leu Thr Leu

1 5 10 15

Ala Ala Gly Ser Thr Pro Ala Glu Thr Ile Ser Asp Ser Ala Glu Ala
20 25 30

Ser Leu Ser Ala Thr Pro Ser Leu Val Thr Trp Thr Gln Val Ser Gly 35 40 45

Leu Gln Pro Leu Val Glu Pro Cys Leu Arg Gln Thr Leu Lys Leu Leu 50 60

Ser Arg Pro Glu Met Trp Arg Ala Val Gly Pro Val Pro Val Ala Cys 65 70 75 80

Leu Leu Phe Leu Gly Ala Tyr Tyr Gln Ala Trp Ser Gln Gln Pro Ser 85 90 95

Ser Cys Pro Glu Asp Trp Leu Gln Asp Met Glu Arg Leu Ser Glu Ser 100 105 110

Cys Cys Cys His Cys Xaa

80

115

<210> 132 <211> 306 <212> PRT <213> Homo sapiens <220> <221> SITE <222> (180) <223> Xaa equals any of the naturally occurring L-amino acids <220> <221> SITE <222> (197) <223> Xaa equals any of the naturally occurring L-amino acids <220> <221> SITE <222> (306) <223> Xaa equals stop translation <400> 132 Met Ser Glu Asp Arg Pro Met Leu Gln Phe Leu Leu His Thr Ser Phe Leu Ser Pro Leu Phe Ile Leu Trp Leu Trp Thr Lys Pro Ile Ala Arg Asp Phe Leu His Gln Pro Pro Phe Gly Glu Thr Arg Phe Ser Leu Leu Ser Asp Ser Ala Phe Asp Ser Gly Arg Leu Trp Leu Leu Val Val Leu Cys Leu Leu Arg Leu Ala Val Thr Arg Pro His Leu Gln Ala Tyr Leu 65 Cys Leu Ala Lys Ala Arg Val Glu Gln Leu Arg Arg Glu Ala Gly Arg 90 Ile Glu Ala Arg Glu Ile Gln Gln Arg Val Val Arg Val Tyr Cys Tyr 100 Val Thr Val Val Ser Leu Gln Tyr Leu Thr Pro Leu Ile Leu Thr Leu 120 Asn Cys Thr Leu Leu Leu Lys Thr Leu Gly Gly Tyr Ser Trp Gly Leu 130 Gly Pro Ala Pro Leu Leu Ser Pro Arg Pro Ile Leu Ser Gln Arg Cys 150 155

Pro His Arg Leu Trp Gly Gly Arg Ser Pro Ala Asp Cys Ser Ala Asp

165 170 175 Cys Arg Gly Xaa Gly Trp Pro Ala Tyr Ser Pro Leu Pro Pro Trp Arg 180 185 Pro Gly Leu Pro Xaa Leu Val Asp Gly Cys Leu Pro Ala Ala Arg Gln 200 Pro Phe Arg Pro Leu Leu Pro Pro Ala Leu Gly Arg Leu Leu Ala Ala 215 Cys Arg Pro Ser Trp Gly Pro Glu Val Cys Ser Trp Gly Ser Gly Thr Leu Ala Cys Pro Leu Cys Leu Arg Pro Arg Val Pro Ser Cys Lys Val 250 Gly Pro Asp Ser Pro Ala Phe Pro Ser Pro Gln Cys Leu Thr Arg Gly 265 Pro Pro Trp Thr Pro Ser Phe Cys Leu Arg Thr Val Ser Pro Gly Pro 275 280 Ser Ser Met Arg Val Pro Arg Pro Leu Ser Pro Lys Arg Met Cys Gln 295 300 Val Xaa 305 <210> 133 <211> 45 <212> PRT <213> Homo sapiens <220> <221> SITE <222> (45) <223> Xaa equals stop translation <400> 133 Met Ser Tyr Ser Leu Phe Leu Ala Leu Leu Ser Phe Ala Ser Ala Ile Leu Phe Val Ala Gly Thr Ile Ala Gly Thr Gly Gly Leu Ser Phe His Gly Ile Ala Thr Ile Phe Val Leu Thr Gly Lys Trp Xaa

<210> 134

<211> 44

<212> PRT

<213> Homo sapiens

```
<220>
<221> SITE
<222> (6)
<223> Xaa equals any of the naturally occurring L-amino acids
<220>
<221> SITE
<222> (44)
<223> Xaa equals stop translation
Met Gly Arg Leu Gly Xaa Gln Cys Leu Leu Phe Leu Ala Phe Lys Ala
                                    10
Ile Ser Gly Val Phe Phe Leu Phe Trp Arg Pro Ala Asp Ser Thr Glu
Arg Asn Thr Gln Ser Trp Asp Phe Pro Pro Leu Xaa
                            40
<210> 135
<211> 50
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (50)
<223> Xaa equals stop translation
<400> 135
Met Gly Val Gly Val Leu Arg Ile Leu Leu Ser Cys Leu Gly Glu Ala
                                    10
Ala Pro Lys Ser Ala Gly Thr Ser Leu Glu Ser Ala Lys Glu Cys Trp
             20
Ser Ala Ala Thr Leu Leu Val Leu Cys Val Leu Cys Gln Leu Gln His
Gly Xaa
<210> 136
<211> 81
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (81)
<223> Xaa equals stop translation
```

83

<400> 136 Met Glu Ser Leu Pro Glu Asn Lys Pro Leu Val Trp Ser Leu Ala Val 10 Ser Leu Leu Ala Ile Ile Gly Leu Leu Gly Ser Ser Pro Asp Phe Asn Ser Gln Phe Gly Leu Val Asp Ile Pro Val Glu Phe Lys Leu Val Ile Ala Gln Val Leu Leu Leu Asp Phe Cys Leu Ala Leu Leu Ala Asp Arg Val Leu Gln Phe Phe Leu Gly Thr Pro Lys Leu Lys Val Pro Ser 75 Xaa <210> 137 <211> 277 <212> PRT <213> Homo sapiens <220> <221> SITE <222> (94) <223> Xaa equals any of the naturally occurring L-amino acids <220> <221> SITE <222> (103) <223> Xaa equals any of the naturally occurring L-amino acids <220> <221> SITE <222> (277) <223> Xaa equals stop translation <400> 137 Met Ile His Val Asn Arg Asn Ile Met Asp Phe Lys Leu Phe Leu Val 10 Phe Val Ala Gly Val Phe Leu Phe Phe Tyr Ala Arg Thr Leu Glu Ser 20 25 Lys Pro Tyr Phe Leu Leu Leu Gly Asn Cys Ala Arg Cys Ser Asn 40 Asp Ile Val Phe Val Leu Leu Val Lys Arg Phe Ile Arg Ser Ile

Ala Pro Phe Gly Ala Leu Met Val Gly Cys Trp Phe Ala Ser Val Tyr

65					70					75					80
Ile	Val	Cys	Gln	Leu 85	Met	Glu	Asp	Leu	Lys 90	Trp	Leu	Trp	Xaa	Glu 95	Asn
Arg	Ile	Tyr	Val 100	Ser	Gly	Xaa	Val	Leu 105	Ile	Val	Gly	Phe	Phe 110	Ser	Phe
Val	Val	Cys 115	Tyr	Lys	His	Gly	Pro 120	Leu	Ala	His	Asp	Arg 125	Ser	Arg	Ser
Leu	Leu 130	Met	Trp	Met	Leu	Arg 135	Leu	Leu	Ser	Leu	Val 140	Leu	Val	Tyr	Ala
Gly 145	Val	Ala	Val	Pro	Gln 150	Phe	Ala	Tyr	Ala	Ala 155	Ile	Ile	Leu	Leu	Met 160
Ser	Ser	Trp	Ser	Leu 165	His	Tyr	Pro	Leu	Arg 170	Ala	Cys	Ser	Tyr	Met 175	Arg
Trp	Lys	Met	Glu 180	Gln	Trp	Phe	Thr	Ser 185	Lys	Glu	Leu	Val	Val 190	Lys	Tyr
Leu	Thr	Glu 195	Asp	Glu	Tyr	Arg	Glu 200	Gln	Ala	Asp	Ala	Glu 205	Thr	Asn	Ser
Ala	Leu 210	Glu	Glu	Leu	Arg	Arg 215	Ala	Cys	Arg	Lys	Pro 220	Asp	Phe	Pro	Ser
Trp 225	Leu	Val	Val	Ser	Arg 230	Leu	His	Thr	Pro	Ser 235	Lys	Phe	Ala	Asp	Phe 240
Val	Leu	Gly	Gly	Ser 245	His	Leu	Ser	Pro	Glu 250	Glu	Ile	Ser	Leu	His 255	Glu
Glu	Gln	Tyr	Gly 260	Leu	Gly	Gly	Ala	Phe 265	Leu	Glu	Glu	Gln	Leu 270	Phe	Asn
Pro	Ser	Thr 275	Ala	Xaa											
<210> 138 <211> 57 <212> PRT <213> Homo sapiens															
<221 <222	<220> <221> SITE <222> (57) <223> Xaa equals stop translation														
<400 Met 1			Thr	Leu 5	Pro	Ala	Arg	Leu	Arg 10	Ala	Gln	Cys	Ile	Ser 15	Ser

85

Leu Leu Phe Leu Leu Met Gly Leu Leu Ala Met Thr Gly Glu Arg Asn 20 25 30

Gln Gly Thr His Tyr Tyr Glu Phe Ser Gly Phe Ile Phe Lys Ser Gln 35 40 45

Met Met Trp Ser Ile Lys Pro Asn Xaa 50 55

<210> 139

<211> 71

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (71)

<223> Xaa equals stop translation

<400> 139

Met Tyr Leu Trp Phe Ser Phe Ser Thr Val Gly Leu Cys Gly Cys Cys 1 5 10 15

Leu Leu Tyr Arg Ala Cys Gly Phe Ile Trp Tyr Leu Leu Leu Gly
20 25 30

His Ser Ser Thr Asn Ser Leu Gln Asp Gly Gly Ala Glu Arg Pro Glu
35 40 45

His Pro Trp Ala His Val Arg Tyr Ser Cys Arg Arg Glu Leu Ser Phe 50 60

Trp Phe Tyr Val Phe Asn Xaa 65 70

<210> 140

<211> 63

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (63)

<223> Xaa equals stop translation

<400> 140

Met Glu Pro Glu Ser Trp Ala Leu Cys Leu Leu Phe Leu Gly Thr

Ala Leu Gly Tyr Pro Pro Leu Pro Arg His Ser Ser Lys Cys Glu Ile 20 25 30

86

Leu Glu Val Arg Leu His Leu Leu Pro Leu Leu Ile Asn Ile Gly Met
35 40 45

Met Ser Pro Val Ala Ser Pro Phe Val Cys Ser Ile Thr Gly Xaa 50 55 60

<210> 141

<211> 89

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (89)

<223> Xaa equals stop translation

<400> 141

Met Leu Phe Leu Ser Ala Ser Ile Cys Thr Ser Ala Leu Phe Leu Cys

1 10 15

Leu Ser Arg Leu Thr Ile Ser Ala Pro His Pro Ala Trp Trp Gly Arg
20 25 30

Met Pro Thr His Thr Ser Pro Gly His Leu Leu Glu Leu Gln Pro Arg 35 40 45

Gly Met Thr Glu Ser Ile Leu Phe Ser Ile Ser Ala Leu Val Ser Asn 50 55 60

Ser Trp Gly Lys Met Thr Gln Leu Thr Ser Gly Ser His Ser Trp Ser 65 70 75 80

Ser Gly Leu Gln Asn Phe Gln Ala Xaa 85

<210> 142

<211> 46

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (46)

<223> Xaa equals stop translation

<400> 142

Met Arg Pro Val Cys Ser Leu Gly Trp Ala Gly Trp Pro Gly Leu Val

Cys Gly Leu Arg Ala Leu Leu Gly Pro Ser Leu Phe Pro Val Thr Phe 20 25 30

Gly Ala Thr Glu Ala Val His Ser Leu Asp Val Cys Ser Xaa

87

35 40 45

<210> 143

<211> 56

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (56)

<223> Xaa equals stop translation

<400> 143

Met Val Asn Glu Lys Glu Ala Arg Thr Gly Ser Pro Lys Ser Trp Leu 1 5 10 15

Leu Cys Leu Ala Leu Leu Leu Ile Lys Tyr Val Thr Phe Cys Lys Pro $20 \hspace{1cm} 25 \hspace{1cm} 30$

Tyr Leu Thr Lys Pro Tyr Phe Leu His Leu Ser Val Leu Asp Gln Leu 35 40

Ser Pro Gly Thr Pro Leu Asp Xaa

<210> 144

<211> 77

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (77)

<223> Xaa equals stop translation

<400> 144

Met Phe Ile Ala Ile Tyr Phe Lys Ala Phe His Gly Ser Phe Gln Leu 1 5 10 15

Cys Thr Trp Leu Val Ile Met Ile Val Ile Leu Gly Gln Ser Phe Ser 20 25 30

Ala Leu Ala Leu Leu Thr Phe Trp Leu Ile Leu Cys Cys Arg Gly Cys 35 40 45

Pro Val His Cys Arg Val Phe Ser Ser Ile Pro Asp Leu Tyr Leu Leu 50 55 60

Asn Ala Arg Ser Asn Thr Val Pro Pro Ala Gln Leu Xaa 65 70 75

```
<211> 43
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (43)
<223> Xaa equals stop translation
<400> 145
Met Phe Phe Leu Ser Met Phe Leu His Ile Val Leu Leu His Cys Gly
                 5
Asn Ser Phe Tyr Lys Ile Cys His Ser Trp Asp Tyr Ala Ala Leu Gln
                                 25
Glu Ser Thr Arg Phe Tyr Ser Asn Ser Tyr Xaa
<210> 146
<211> 102
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (67)
<223> Xaa equals any of the naturally occurring L-amino acids
<220>
<221> SITE
<222> (102)
<223> Xaa equals stop translation
<400> 146
Met Glu Leu Glu Arg Cys Ser Val Val Leu Cys Ile Leu Ala Asn Leu
Ala Val Leu Arg Ala Leu Phe Leu Pro Cys Ile Ile Phe His Cys Val
                                 25
Ser Asp Ser Arg Ser Val Asn Arg Glu Thr Lys Val Lys Phe Val His
                             40
                                                 45
         35
Thr Ser Val His Gly Val Gly His Ser Phe Val Gln Ser Ala Phe Lys
Ala Phe Xaa Leu Val Pro Pro Glu Ala Val Pro Glu Gln Lys Asp Pro
Asp Pro Glu Phe Pro Thr Val Lys Tyr Pro Asn Pro Glu Glu Gly Lys
```

Gly Val Leu Val Thr Xaa

89

100

<223> Xaa equals stop translation

5

<400> 148

```
<210> 147
<211> 134
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (134)
<223> Xaa equals stop translation
<400> 147
Met Arg Val Pro Leu Val Leu Ser Trp Ala Phe Val Leu Val Gly Phe
                                    10
Ser Gly Val Tyr Leu Ala Ser Glu Ser Phe Trp Phe Pro Pro Ser Leu
                                 25
Cys Asp Leu Thr Ser Pro Pro Gly Leu His Leu Trp Lys Phe Ile Arg
        35
                            40
Asp Leu Val Ser Met Glu Glu Leu Thr Asp Ser Ala Arg Glu Met Gly
Tyr Trp Met Met Val Phe Ser Leu Lys Ala Met Phe Pro Val Ser Ser
                     70
Gly Cys Phe Gln Glu Arg Gln Glu Thr Asn Lys Ser Leu Thr Leu Leu
Arg Cys Ser Gln Arg Asp Thr Ser Pro Leu Met Asp Gly Gln Thr Trp
                105
Ala Arg Val Arg Val Thr Lys Pro Pro Thr Thr Ala Thr Ala Ala Tyr
                           120
Asn Arg His Ile Arg Xaa
   130
<210> 148
<211> 50
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (50)
```

Met Lys Ser Leu Phe Cys Ile Tyr Phe Leu Arg Trp Pro Met Gly Leu

90

Ser Trp Gly Glu Thr Phe Ile Leu Leu Arg Asp Ser Leu Ala Ile Asn $20 \hspace{1cm} 25 \hspace{1cm} 30$

Phe Gln Ser Phe Ser Lys Ala Ala Ser Gly Asp Ile Phe Gly Cys His
35 40 45

Asp Xaa 50

<210> 149

<211> 64

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (6)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (64)

<223> Xaa equals stop translation

<400> 149

Met Ser Cys Gly Leu Xaa Phe Gly Pro Trp Phe Val Pro Met Leu Leu 1 5 10 15

Met Ser His Ser Leu Leu Pro Ser Trp Ser Gly Leu Trp Val Thr Thr 20 25 30

Trp Asn Gly Ser Ser Gly Glu Arg Thr Pro Ser Pro Trp Arg Arg Lys
35 40 45

Arg Ala Ser Gln Ser Ala Gly Arg Ile Ala Ser Trp Met Ser Phe Xaa 50 55 60

<210> 150

<211> 75

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (59)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

PCT/US99/05804

91

```
<222> (75)
<223> Xaa equals stop translation
```

<400> 150

WO 99/47540

Met Leu Ser Ser Pro Asn Leu Ala Ala Ser Leu Leu Cys Leu Trp His 1 5 10 15

Ser Gly Ser Ala Thr Asn Trp Ala Pro Pro Cys Ala Gly Met Trp Ala 20 25 30

Ser Arg Cys Gly Trp Lys Val Ser Pro His Pro Glu Ala Gly Pro Cys 35 40

Ser Ser Ala Leu Trp Val Ser Cys Cys Val Xaa Ala Glu Gln Pro Gln 50 55 60

Pro Gly Gly Arg Glu Pro Arg His Arg Gly Xaa 65 70 75

<210> 151

<211> 55

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (55)

<223> Xaa equals stop translation

<400> 151

Met Pro His Ile Ser Phe Cys Leu Gly Thr Pro Tyr Val Val Ala Val 1 5 10 15

Tyr Leu Pro Ala Trp Ile Val Met Leu Leu Pro Gly Val Arg Pro 20 25 30

Tyr Ser Ser Leu Gln Ala Leu Lys His Pro Ser Cys Ser Ser Ser Ser Ser 35 40 45

Val Cys Ala Pro Tyr Met Xaa 50 55

<210> 152

<211> 58

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (58)

<223> Xaa equals stop translation

<400> 152

92

Met Gly Leu Asn Ile Ser Pro Trp Cys Phe Leu Ala Ile Leu Thr Cys

Ala Ile Ser Ala Ala Phe Ile Ser Val Gly Val Val Cys Trp Leu Leu

Phe Leu Ile Ser His Arg Ser Ser Lys Asn Leu Arg Lýs Ser Arg Val

Arg Gly Val Trp Glu Asn Glu Glu Ile Xaa

<210> 153

<211> 53

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (53)

<223> Xaa equals stop translation

<400> 153

Met Ala Tyr Val Leu Ala Val Leu Cys Phe Lys Ser Leu Trp Ala Leu

Phe Lys Pro Asn Lys Gln Leu Ile Glu Phe Leu Leu Met Val Lys Val 25

Val Lys Ile Pro Leu Cys Tyr Leu Arg Gln Leu Leu Gly Gly Ile Lys 35 40

Thr Pro Arg Val Xaa 50

<210> 154

<211> 51

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (51)

<223> Xaa equals stop translation

<400> 154

Met Asp Gly Gly Pro Gly Ala Phe Ser Arg Ala Trp Val Leu Gln Ile

Pro Trp Leu Leu Ser Gly Gly Asn Phe Ala Leu Cys Glu Pro Arg 25

Pro Cys Pro Ser Ala Gly His Pro Trp Gln Glu Ala Gly Leu Pro Ser

PCT/US99/05804

93

35 40 45 Ser Pro Xaa 50 <210> 155 <211> 67 <212> PRT <213> Homo sapiens <220> <221> SITE <222> (55) <223> Xaa equals any of the naturally occurring L-amino acids <220> <221> SITE <222> (67) <223> Xaa equals stop translation <400> 155 Met Pro Phe Leu Ser Val Trp Phe Phe Asn Leu Gly Leu Ile Phe Gly Val Glu Ser Phe Val Leu Arg Ala Val Leu Phe Ile Ala Gly Cys Ser Ala Thr Ser Gln Met Glu Ala Ala Ser Pro Tyr Pro Ala Val Thr Lys Arg Lys Lys Asn Val Ser Xaa His Cys Gln Ile Ser Ser Gly Gly Ala Pro Gly Xaa 65 <210> 156 <211> 49 <212> PRT <213> Homo sapiens <220> <221> SITE <222> (49) <223> Xaa equals stop translation <400> 156 Met Leu Leu Lys Arg Asn Leu Leu Ile Leu Ile Leu Phe Leu Val Thr Cys Phe Asn Phe Val Ser Phe Phe Phe Phe Pro Trp Lys Leu Leu Gly

25

Ser Pro Phe Tyr Pro Cys Ser Leu Arg Ser Asp Asn Asp Gly Cys Val 35 40 45

Xaa

<210> 157

<211> 61

<212> PRT

<213> Homo sapiens

<400> 157

Met Gly Ser Phe Leu His Pro Gln Trp His Leu Leu Ile Thr Phe Cys

1 10 15

Ala Val Leu Gly Lys Gly Leu His Ser Asp Pro Ser Arg Pro Phe Glu 20 25 30

His Gly Gly Ala Leu Gly Lys Val Pro Arg Gly Arg Ser Thr Leu Leu 35 40 45

Ser Lys Glu Val Leu Leu Lys Lys Lys Lys Lys Lys Arg

<210> 158

<211> 118

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (113)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (118)

<223> Xaa equals stop translation

<400> 158

Met Leu Leu Trp Trp Gln Cys Leu Cys Cys His Ala Val Leu Glu Pro 1 5 10 15

Ala Ala Thr Ala Met Pro Glu Asp Ala Ala Pro Ser Ser Leu Pro Val 20 25 30

Pro Pro Asn Met Thr Ser Ser Arg Phe His Tyr Phe Trp Thr Leu Leu 35 40

Gln Ile Lys Leu Thr Gln Phe Tyr Ser Lys Pro Arg Ser Leu Ser Ala 50 55 60

Thr Pro Glu Lys Asn Ile Gly Leu Gln Glu Pro Glu Arg Arg Glu Arg

70 75 80 65 Phe Thr Gly Glu Ser Cys Arg Trp Glu Leu Lys Ala Lys Ser Cys Leu 90 Cys Pro Thr Arg Asn Ser Leu Gly Cys Thr Gln Cys His Cys Asp Gly 105 Xaa Lys Ile Cys Asn Xaa 115 <210> 159 <211> 151 <212> PRT <213> Homo sapiens <220> <221> SITE <222> (151) <223> Xaa equals stop translation <400> 159 Met Leu Ala Val Leu Ala Phe Pro Val Gly Val Phe Val Val Ala Val Phe Trp Ile Ile Tyr Ala Tyr Asp Arg Glu Met Ile Tyr Pro Lys Leu 25 Leu Asp Asn Phe Ile Pro Gly Trp Leu Asn His Gly Met His Thr Thr Val Leu Pro Phe Ile Leu Ile Glu Met Arg Thr Ser His His Gln Tyr Pro Ser Arg Ser Ser Gly Leu Thr Ala Ile Cys Thr Phe Ser Val Gly 70 Tyr Ile Leu Trp Val Cys Trp Val His His Val Thr Gly Met Trp Val Tyr Pro Phe Leu Glu His Ile Gly Pro Gly Ala Arg Ile Ile Phe Phe 105 100 Gly Ser Thr Thr Ile Leu Met Asn Phe Leu Tyr Leu Leu Gly Glu Val

Leu Asn Asn Tyr Ile Trp Asp Thr Gln Lys Ser Met Glu Glu Glu Lys
130 135 140

120

Glu Lys Pro Lys Leu Glu Xaa 145 150

```
<211> 92
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (92)
<223> Xaa equals stop translation
<400> 160
Met Gly Asp Lys Leu Gly Met Ala Arg Ala Pro Ser Val Ala Leu Ala
Gln Leu Trp Leu Ile Cys Leu Cys Pro Glu Ser Leu Ala Ser Phe Val
Gln Ala Val Pro Trp Lys Val Leu Gln Pro Ser Ser Asn Arg Ser Thr
Asp Cys Ser Pro His Met Arg Pro Thr Cys Glu Thr Leu Gly Ser Arg
Lys Ala Gln Asp Leu Val Leu Asp Thr Met Cys Leu Ser Thr Asp Asp
Cys Gln Gly Leu Ile Cys Arg Gly His Arg Ser Xaa
                 85
<210> 161
<211> 42
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (42)
<223> Xaa equals stop translation
<400> 161
Met Gln Val Ala Cys Val Met Lys Val Ser Ala Gln Trp Val Cys Phe
Phe Val Val Phe Ser Pro Leu Cys Ser Ser Val Lys Cys Ala Ser Ser
Gly Gln Asn Arg Gly Arg Gly Asp Gln Xaa
                             40
         35
```

<210> 162 <211> 78 <212> PRT <213> Homo sapiens

```
<220>
<221> SITE
<222> (78)
<223> Xaa equals stop translation
Met Met Leu Gln Ile Ile His Leu Asn Thr Leu Ile Lys Phe Phe Gln
                                  10
                5
Cys Leu Lys Leu Phe Leu His Gly Thr Ala Gly Ser Gly Gln Lys Cys
                               25
Leu Ala Tyr Lys Phe Ser Gln Phe Pro Ser Ile Ile Pro Ala Ala His
                                             45
Lys Lys Val His His Leu Leu Ser Pro Lys Cys Leu Pro Thr Glu Cys
Ser Gln Ala Asp Asn Ser Ser Trp Asp Ser Ala Val Trp Xaa
65 70
<210> 163
<211> 55
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (55)
<223> Xaa equals stop translation
<400> 163
Met Lys Arg Leu Trp Cys Leu Ser Trp Val Pro Gly Leu Gln Gly Ser
Pro Ser Val Leu Ser Ser Val Phe Phe Ser Val Phe Lys Pro Gln Leu
        20 25 30
His Trp Thr Cys Ser Gln Val Ser Ser His Trp His Pro Pro Cys Leu
Phe Ile Leu Phe Ser Gly Xaa
<210> 164
<211> 90
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (90)
```

<223> Xaa equals stop translation

```
<400> 164
Met Lys Phe Leu Leu Ala Ala Leu Val Leu Ser Leu Ile Leu Pro Arg
Ser Ser Gln Tyr Ile Lys Trp Ile Val Ser Ala Gly Leu Ala Gln Val
                               25
Ser Glu Phe Ser Phe Val Leu Gly Ser Arg Ala Arg Arg Ala Gly Val
Ile Ser Arg Glu Val Tyr Leu Leu Ile Leu Ser Val Thr Thr Leu Ser
Leu Leu Ala Pro Val Leu Trp Arg Ala Ala Ile Thr Arg Cys Val
Pro Arg Pro Glu Arg Arg Ser Ser Leu Xaa
                85
<210> 165
<211> 45
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (45)
<223> Xaa equals stop translation
<400> 165
Met Phe Val Trp His Leu Lys Val Met Val Met Phe Ile Ile Leu Tyr
Phe Ala Tyr Cys Glu Ser Asn Phe His Ser Val Leu Ser Val Ser Lys
                               25
Pro Leu Leu Lys Ile Leu Phe Leu Pro Arg Asn Leu Xaa
               40
<210> 166
<211> 45
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (45)
<223> Xaa equals stop translation
Met Thr Pro Gly Cys Ser Val Pro Phe Leu Leu Cys Trp Leu Phe Ala
                 5
                                    10
```

```
Leu Met Met Gln Glu Lys Trp Gly Gly Val Lys Ser Leu Val Ser Tyr
                                 25
```

His Tyr Ser Arg Gln Trp His Gln Thr Val Val Val Xaa 35 40

<210> 167

<211> 66

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (66)

<223> Xaa equals stop translation

<400> 167

Met Ser Ile Ala Leu Arg Ile Asn Arg Leu His Phe Trp Val Leu Leu

Phe Phe Phe Phe Ala Gln Leu Ser Leu Ser Val Asp Leu His Gly 20

Thr Ser Tyr Ser Leu Lys Ser Leu Ser Tyr Leu Thr Ile Phe Leu Asp

Leu Glu Lys Leu Asp Val Gly Pro Tyr Glu Lys Ile Ile Arg Asn Gln

Ile Xaa

65

<210> 168

<211> 62

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (62)

<223> Xaa equals stop translation

<400> 168

Met Gln Leu Thr Leu Gly Gly Ala Ala Val Gly Ala Gly Ala Val Leu 10

Ala Ala Ser Leu Leu Trp Ala Cys Ala Val Gly Leu Tyr Met Gly Gln 20

Leu Glu Leu Asp Val Glu Leu Val Pro Glu Asp Asp Gly Thr Ala Ser 40 45

```
Ala Glu Gly Pro Asp Glu Ala Gly Arg Pro Pro Pro Glu Xaa
                         55
<210> 169
<211> 47
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (47)
<223> Xaa equals stop translation
<400> 169
Met His Thr Ala Lys Met Ser Leu Leu Asn Ser Val Cys Leu Leu Val
Leu Ser Ile Trp Tyr Val Val Lys Phe Pro Met Met Arg Asp Ser Thr
Ile Asn Val Pro Tyr Leu Leu Arg Leu Lys Ala Ile Thr Thr Xaa
                             40
<210> 170
<211> 106
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (69)
<223> Xaa equals any of the naturally occurring L-amino acids
<220>
<221> SITE
<222> (1.06)
<223> Xaa equals stop translation
Met Ser Gly Leu Ala Ala Ala His Val Phe Arg Val Cys Leu Phe
Pro Leu Ser Trp Gly Ser Ser Lys Thr Thr Phe Ile His Gly Leu Ser
                                                     30
             20
                                 25
Ser Tyr Ile Ala Thr Pro Val Leu Asn Ser Ile Phe Ser Ser Trp Lys
Ser Arg Arg Lys Asp Thr Trp Thr Cys Leu Leu His Arg Leu Ser Ala
                         55
     50
Phe Pro Ile Ser Xaa Arg Arg Arg Asn Phe Ala Leu Phe Ser His Ser
65
                     70
                                         75
```

Cys Val Cys Ile Arg Ser Ser Ser Asp Asp Val Gly Pro Thr Met Tyr 85 90 Ser Phe Ser Val Pro Cys Arg Val Lys Xaa <210> 171 <211> 45 <212> PRT <213> Homo sapiens <220> <221> SITE <222> (45) <223> Xaa equals stop translation <400> 171 Met His Leu Leu Thr Leu Phe Ser Ser Gly Leu Ile Phe Leu Gly Cys Ser Thr Pro Leu Ser Phe Cys Asp Cys Leu Pro Ile Leu Leu Leu Trp 25 Leu Glu Phe Pro Val Glu Thr Ser Gly Val Cys Ser Xaa 40 <210> 172 <211> 47 <212> PRT <213> Homo sapiens <220> <221> SITE <222> (47) <223> Xaa equals stop translation <400> 172 Met Ile Leu Lys His Tyr Ile Leu Thr Phe Ile Phe Leu Phe Ile Phe 1 10 Leu Phe Phe Met Leu Asn Ile Leu His Ser Asn Ser Asn Leu Ile Asp 25 Leu Leu Lys Gly Asn Ile Arg Phe Arg Leu Leu Asn Ser Met Xaa

<210> 173 <211> 42 <212> PRT <213> Homo sapiens

PCT/US99/05804 WO 99/47540

102

<220> <221> SITE <222> (42) <223> Xaa equals stop translation

<400> 173 Met Ala Thr Leu Gln Ile Thr Thr Ala Met Lys Ile Thr Met Met Ile

Thr Met Val Met Ile Ile Thr Thr Ile Val Glu Ala Met Lys Ile Pro

Thr Thr Ala Met Met Met Ala Met Gln Xaa

<210> 174 <211> 47

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (47)

<223> Xaa equals stop translation

<400> 174

Met Glu Met Leu Ser Ser Lys Trp Ser Lys Arg Val Ala Ala Ser Leu 10

Ala His Leu Ile Ser Leu Phe Ile Gly Leu Leu Phe Leu Leu Gly 25 20

Ser Ser Val Tyr Pro Gly Thr Glu Thr Leu Phe Pro Lys Ser Xaa 40

<210> 175

<211> 41

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (41)

<223> Xaa equals stop translation

<400> 175

Met Trp Pro Ser Leu Gly Arg Cys Cys Leu Phe Phe Cys Leu Leu Thr 10

Asn Leu Thr Ser Cys His Thr Ser Gln Ile Thr Leu Cys Ser Arg Glu

Thr Cys Val Trp Ser Arg Thr Thr Xaa

103

40

<210> 176 <211> 53 <212> PRT <213> Homo sapiens <220>

35

<221> SITE <222> (53) <223> Xaa equals stop translation

Cys Thr Ile Leu Val Leu Ser Pro Pro Val Leu Leu Lys Tyr Gln Asp 20 25 30

Ser Thr Pro Arg Pro Leu Trp Ser Gln Cys Lys Ile Pro Ile Asn Tyr 35 40 45

Leu Lys Gly Lys Xaa 50

<210> 177 <211> 250

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (250)

<223> Xaa equals stop translation

<400> 177

Met Arg Gly Pro Ser Trp Ser Arg Pro Arg Pro Leu Leu Leu Leu 1 5 10 15

Leu Leu Ser Pro Trp Pro Val Trp Ala Gln Val Ser Ala Arg Ala
20 25 30

Ser Pro Ser Gly Ser Leu Gly Ala Pro Asp Cys Pro Glu Val Cys Thr

Cys Val Pro Gly Gly Leu Pro Ala Val Gly Thr Leu Ala Ala Arg Arg 50 55 60

Ala Pro Gly Pro Glu Pro Ala Pro Ala Arg Ala Ala Gly Pro Gln
65 70 75 80

Pro Arg Pro Cys Ala Ala Ala Arg Cys Leu Arg Gly Ser Gly Arg Ala 85 90 95

Thr Ala Pro Gly Pro Ala Arg Glu Arg Ala Ala Leu Gly Ala Cys Ala 100

Ser Leu Leu Gly Pro Gly Arg Ala Ala Ala Ala Gly Pro Glu Arg Gln 115

Pro Ala Gly Ser Thr Gly Thr Arg Asp Phe Arg Ala Ala Ala Ala Arg Ala 130

Ala Gln Pro Leu Ile Gly Arg Gln Pro Ala Gly Ala Pro Gly Ala Arg 145 150 155 160

Gly Ala Arg Arg Ala Pro Ala Ala Ala Leu Thr Gln Pro Ala Gly Gln 165 170 175

Arg Ala Gly Gly Thr Arg Ala Gly Ala Ala Gly Pro Pro Ala Arg Ser 180 185 190

Arg Arg Ala Ala Pro Ala Arg Gln Pro Leu Gly Leu Arg Val Arg Ala 195 200 205

Ala Pro Ala Leu Arg Leu Ala Ala Pro Ala Pro Ala Ala Arg Val Arg 210 215 220

Gly Arg Asp Gly Ala Leu Arg Val Ala Gly Thr Pro Asp Ala Gln Pro 225 230 235 240

Pro Asp Cys Leu Phe Arg Arg Arg Leu Xaa 245

<210> 178

<211> 148

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (148)

<223> Xaa equals stop translation

<400> 178

Met Leu Ala Gly Ala Gly Arg Pro Gly Leu Pro Gln Gly Arg His Leu

1 5 10 15

Cys Trp Leu Leu Cys Ala Phe Thr Leu Lys Leu Cys Gl
n Ala Glu Ala 20 \$25\$

Pro Val Glu Glu Lys Leu Ser Ala Ser Thr Ser Asn Leu Pro Cys
35 40 45

Trp Leu Val Glu Glu Phe Val Val Ala Glu Glu Cys Ser Pro Cys Ser 50 60

```
Asn Phe Arg Ala Lys Thr Thr Pro Glu Cys Gly Pro Thr Gly Tyr Val 65 70 75 80
```

Glu Lys Ile Thr Cys Ser Ser Ser Lys Arg Asn Glu Phe Lys Ser Cys 85 90 95

Arg Phe Ser Phe Glu Trp Asn Asn Ala Tyr Phe Gly Ser Ser Lys Gly 100 105 110

Ala Val Val Cys Val Ala Leu Ile Phe Ala Cys Leu Val Ile Ile Arg 115 120 125

Gln Arg Gln Leu Asp Arg Lys Ala Leu Glu Lys Val Arg Lys Gln Ile 130 135 140

Glu Ser Ile Xaa 145

<210> 179

<211> 48

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (48)

<223> Xaa equals stop translation

<400> 179

Met Phe Met Cys Arg Leu Leu Leu Trp Ala Thr Gly Ala Tyr Gly Phe 1 5 10 15

Leu Gly Asp Asp Val Glu Tyr Thr Ser Val Leu Pro His Gln Lys Gly 20 25 30

Lys Glu Ala Trp Val Phe Ile Cys Gln Leu Pro Phe Ile Ile Gly Xaa 35 40 45

<210> 180

<211> 57

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (56)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

106

```
<222> (57)
<223> Xaa equals stop translation
Met Leu Gln Thr Leu Leu Cys Leu Trp Gln Tyr Thr Ser Ala Gln Val
                                 10
Leu Lys Met Leu Cys Ile His Arg Gln Lys Trp Asp Asn Phe Trp Ala
                                25
Val Val Met Ile Asn Leu Leu Ile Arg Ile Gln Arg Leu Pro Phe Ser
                            40
Leu Pro Ile Ala Leu Arg Val Xaa Xaa
                        55
<210> 181
<211> 49
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (49)
<223> Xaa equals stop translation
<400> 181
Met Pro Ser Glu Gly Arg Leu Val Leu Leu Ser Ala Phe Cys Pro Ala
Phe Phe Pro Pro Trp Val Leu Ser Gly Ser Phe Ala Phe Ser Leu Cys
                               25
Ala Glu Ser His Leu Asn Ser Ser His Arg Arg Ile Ala Val Trp Thr
                           40
Xaa
<210> 182
<211> 46
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (46)
<223> Xaa equals stop translation
<400> 182
Met Val Gln Trp Lys Asn Trp Pro Glu Ser Leu Glu Val Trp Val Leu
```

107

Val Leu Ala Val Pro Leu Thr His Cys Asp Leu Gly Ile Leu Cys Cys 20 25 30

Glu Asp Ile Ser Gln Val Leu His Val Ser Gln Gln Ile Xaa 35 40 45

<210> 183

<211> 41

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (41)

<223> Xaa equals stop translation

<400> 183

Met Ala Leu Gly Leu Cys Ser Ser Gly Ala Leu Ser Thr Leu Cys Leu 1 5 10 15

Ser Ser Val Thr Cys Leu Ala Ile Met Val Leu Met Ala Val Asp Gly 20 25 30

Leu His Gly Thr Ser Gly Leu Gly Xaa 35 40

<210> 184

<211> 80

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (80)

<223> Xaa equals stop translation

<400> 184

Met Thr Leu Met Cys Leu Cys Leu Ser Val Thr Val Leu His Pro Leu $1 \hspace{1cm} 5 \hspace{1cm} 10 \hspace{1cm} 15$

Arg Ser Lys Glu Arg Leu Ser Gly Thr Phe Cys Gly Tyr Ser Ser Ser 20 25 30

Trp Cys Ser Pro Ala Ser Glu Ser Ser Pro Gly Ser Leu Leu Thr 35 40 45

Cys Ala Ala Ser Gly Ser His Pro Asp Cys Pro Leu Ser Gln Arg Leu 50 60

Leu Gly Val Gln Leu Ala Ala Leu Gly Arg Pro Gln Gly Leu Phe Xaa 65 70 75 80

108

```
<210> 185
<211> 47
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (47)
<223> Xaa equals stop translation
Met Lys Ser Gln Cys Tyr Ser Pro Ser Tyr Phe Ala Phe Phe Cys Leu
                                     10
Val Phe Phe Gln Ile Thr Ser Ala Ser Ser Gln Thr Leu Arg Gly His
                                25
Val Leu Cys Arg Thr Thr Leu Arg Asp Ser Ser Ala Tyr Cys Xaa
         35
                            40
<210> 186
<211> 141
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (36)
<223> Xaa equals any of the naturally occurring L-amino acids
<220>
<221> SITE
<222> (141)
<223> Xaa equals stop translation
<400> 186
Met Phe Leu Phe Gly Gly Phe Leu Met Thr Leu Phe Gly Leu Phe Val
Ser Leu Val Phe Leu Gly Gln Ala Phe Thr Ile Met Leu Val Tyr Val
Trp Ser Arg Xaa Asn Pro Tyr Val Arg Met Asn Phe Phe Gly Leu Leu
Asn Phe Gln Ala Pro Phe Leu Pro Trp Val Leu Met Gly Phe Ser Leu
Leu Leu Gly Asn Ser Ile Ile Val Asp Leu Leu Gly Ile Ala Val Gly
                                        75
```

109

His Ile Tyr Phe Phe Leu Glu Asp Val Phe Pro Asn Gln Pro Gly Gly 90 Ile Arg Ile Leu Lys Thr Pro Ser Ile Leu Lys Ala Ile Phe Asp Thr 105 Pro Asp Glu Asp Pro Asn Tyr Asn Pro Leu Pro Glu Glu Arg Pro Gly Gly Phe Ala Trp Gly Glu Gly Gln Arg Leu Gly Gly Xaa 135 <210> 187 <211> 339 <212> PRT <213> Homo sapiens <220> <221> SITE <222> (339) <223> Xaa equals stop translation <400> 187 Met Arg Lys Pro Ala Ala Gly Phe Leu Pro Ser Leu Leu Lys Val Leu 10 Leu Leu Pro Leu Ala Pro Ala Ala Gln Asp Ser Thr Gln Ala Ser Thr Pro Gly Ser Pro Leu Ser Pro Thr Glu Tyr Glu Arg Phe Phe Ala 35 45 Leu Leu Thr Pro Thr Trp Lys Ala Glu Thr Thr Cys Arg Leu Arg Ala 55 Thr His Gly Cys Arg Asn Pro Thr Leu Val Gln Leu Asp Gln Tyr Glu 70 75 Asn His Gly Leu Val Pro Asp Gly Ala Val Cys Ser Asn Leu Pro Tyr Ala Ser Trp Phe Glu Ser Phe Cys Gln Phe Thr His Tyr Arg Cys Ser Asn His Val Tyr Tyr Ala Lys Arg Val Leu Cys Ser Gln Pro Val Ser Ile Leu Ser Pro Asn Thr Leu Lys Glu Ile Glu Ala Ser Ala Glu Val 135 Ser Pro Thr Thr Met Thr Ser Pro Ile Ser Pro His Phe Thr Val Thr

Glu Arg Gln Thr Phe Gln Pro Trp Pro Glu Arg Leu Ser Asn Asn Val

110

165 170 175 Glu Glu Leu Leu Gln Ser Ser Leu Ser Leu Gly Ser Gln Glu Gln Ala 185 Pro Glu His Lys Gln Glu Gln Gly Val Glu His Arg Gln Glu Pro Thr Gln Glu His Lys Gln Glu Glu Gly Gln Lys Gln Glu Glu Glu Glu Glu 215 Glu Gln Glu Glu Gly Lys Gln Glu Glu Gly Gln Gly Thr Lys Glu 230 235 Gly Arg Glu Ala Val Ser Gln Leu Gln Thr Asp Ser Glu Pro Lys Phe 245 250 His Ser Glu Ser Leu Ser Ser Asn Pro Ser Ser Phe Ala Pro Arg Val 265 Arg Glu Val Glu Ser Thr Pro Met Ile Met Glu Asn Ile Gln Glu Leu 275 280 Ile Arg Ser Ala Gln Glu Ile Asp Glu Met Asn Glu Ile Tyr Asp Glu 295 Asn Ser Tyr Trp Arg Asn Gln Asn Pro Gly Ser Leu Leu Gln Leu Pro 305 310 315 His Thr Glu Pro Cys Trp Cys Cys Ala Ile Arg Ser Trp Arg Ile Pro 330 Ala Ser Xaa <210> 188 <211> 66 <212> PRT <213> Homo sapiens <220> <221> SITE <222> (66) <223> Xaa equals stop translation <400> 188 Met Gln Arg Ile Pro Thr Ser Pro Arg Gln Ala Trp Trp Trp Thr Cys Trp Ala Met Phe Gln Gly Pro Ala Ala Gly Ser Val Gly Ala Glu Arg Lys Gly Glu Gly Cys Leu Phe Phe Gly Gln Asp Glu Ser Ser Arg Cys

40

111

Gly Arg Ser Trp Pro Leu Ala Asp Pro Trp Val Tyr Arg Val Leu Arg 55 Ser Xaa 65 <210> 189 <211> 360 <212> PRT <213> Homo sapiens <400> 189 Met Val Pro Ala Ala Gly Arg Arg Pro Pro Arg Val Met Arg Leu Leu 10 Gly Trp Trp Gln Val Leu Leu Trp Val Leu Gly Leu Pro Val Arg Gly 25 Val Glu Val Ala Glu Glu Ser Gly Arg Leu Trp Ser Glu Glu Gln Pro 40 Ala His Pro Leu Gln Val Gly Ala Val Tyr Leu Gly Glu Glu Leu Leu His Asp Pro Met Gly Gln Asp Arg Ala Ala Glu Glu Ala Asn Ala 70 75 Val Leu Gly Leu Asp Thr Gln Gly Asp His Met Val Met Leu Ser Val Ile Pro Gly Glu Ala Glu Asp Lys Val Ser Ser Glu Pro Ser Gly Val 100 105 Thr Cys Gly Ala Gly Gly Ala Glu Asp Ser Arg Cys Asn Val Arg Glu Ser Leu Phe Ser Leu Asp Gly Ala Gly Ala His Phe Pro Asp Arg Glu 135 Glu Glu Tyr Tyr Thr Glu Pro Glu Val Ala Glu Ser Asp Ala Ala Pro Thr Glu Asp Ser Asn Asn Thr Glu Ser Leu Lys Ser Pro Lys Val Asn 170 Cys Glu Glu Arg Asn Ile Thr Gly Leu Glu Asn Phe Thr Leu Lys Ile Leu Asn Met Ser Gln Asp Leu Met Asp Phe Leu Asn Pro Asn Gly Ser 200

Asp Cys Thr Leu Val Leu Phe Tyr Thr Pro Trp Cys Arg Phe Ser Ala

220

215

210

Ser Leu Ala Pro His Phe Asn Ser Leu Pro Arg Ala Phe Pro Ala Leu 225 235 His Phe Leu Ala Leu Asp Ala Ser Gln His Ser Ser Leu Ser Thr Arg Phe Gly Thr Val Ala Val Pro Asn Ile Leu Leu Phe Gln Gly Ala Lys 265 Pro Met Ala Arg Phe Asn His Thr Asp Arg Thr Leu Glu Thr Leu Lys 280 Ile Phe Ile Phe Asn Gln Thr Gly Ile Glu Ala Lys Lys Asn Val Val 295 300 Val Thr Gln Ala Asp Gln Ile Gly Pro Leu Pro Ser Thr Leu Ile Lys 310 315 Ser Val Asp Trp Leu Leu Val Phe Ser Leu Phe Phe Leu Ile Ser Phe 325 330 Ile Met Tyr Ala Thr Ile Arg Thr Glu Ser Ile Arg Trp Leu Ile Pro 345 340 Gly Gln Glu Gln Glu His Val Glu 355 <210> 190 <211> 160 <212> PRT <213> Homo sapiens <220> <221> SITE <222> (160) <223> Xaa equals stop translation <400> 190 Met Leu Leu Leu Ile Phe Trp Ile Ala Pro Ala His Gly Pro Thr Asn Ile Met Val Tyr Ile Ser Ile Cys Ser Leu Leu Gly Ser Phe Thr Val Pro Ser Thr Lys Gly Ile Gly Leu Ala Ala Gln Asp Ile Leu His Asn Asn Pro Ser Ser Gln Arg Ala Leu Cys Leu Cys Leu Val Leu Leu Ala Val Leu Gly Cys Ser Ile Ile Val Gln Phe Arg Tyr Ile Asn Lys 70 75

113

Ala Leu Glu Cys Phe Asp Ser Ser Val Phe Gly Ala Ile Tyr Tyr Val 85 90 95

Val Phe Thr Thr Leu Val Leu Leu Ala Ser Ala Ile Leu Phe Arg Glu 100 105 110

Trp Ser Asn Val Gly Leu Val Asp Phe Leu Gly Met Ala Cys Gly Phe
115 120 125

Thr Thr Val Ser Val Gly Ile Val Leu Ile Gln Val Phe Lys Glu Phe 130 135 140

Asn Phe Asn Leu Gly Glu Met Asn Lys Ser Asn Met Lys Thr Asp Xaa 145 150 155 160

<210> 191

<211> 101

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (92)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (96)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (101)

<223> Xaa equals stop translation

<400> 191

Met Phe Val Ala Val Phe Tyr Trp Val Leu Thr Val Phe Phe Leu Ile 1 5 10 15

Ile Tyr Ile Thr Met Thr Tyr Thr Arg Ile Pro Gln Val Pro Trp Thr 20 25 30

Thr Val Gly Leu Cys Phe Asn Gly Ser Ala Phe Val Leu Tyr Leu Ser 35 40

Ala Ala Val Val Asp Ala Ser Ser Val Ser Pro Glu Lys Asp Ser His
50 60

Asn Phe Asn Ser Trp Ala Ala Ser Ser Phe Phe Ala Phe Leu Val Thr 65 70 75 80

114

Ile Cys Tyr Ala Gly Asn Thr Tyr Phe Ser Phe Xaa Ala Trp Arg Xaa 85 90 95

Arg Thr Ile Gln Xaa 100

<210> 192

<211> 43

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (43)

<223> Xaa equals stop translation

<400> 192

Met Phe Lys Leu Gln Leu Asp Leu Leu Thr Ala Val Asn Leu Val Tyr

1 5 10 15

Phe Ser Phe Leu Trp Val Val Ser Val Ala Asn Lys Met Asp Val Ser 20 25 30

Val Phe Glu Leu Val Asn Ser Asp Cys Phe Xaa 35 40

<210> 193

<211> 62

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (62)

<223> Xaa equals stop translation

<400> 193

Met Ser Val Cys Val Phe Leu Asp Phe Arg Leu Ile Phe Trp Ser Phe 1 5 10 15

Cys Pro Cys Ser Ala Ser Pro Ser Arg His Phe Ala Ser Ser Ser Arg 20 25 30

Gly Gly Gly Gly Ser Arg Asn Trp Val Gly Ala Gly Ala Ser Leu 35 40 45

Ala Ala Ser Leu Ala Leu Tyr Ala Leu Ser Pro Arg Arg Xaa 50 55 60

<210> 194

<211> 53

<212> PRT

115

```
<213> Homo sapiens
<220>
<221> SITE
<222> (53)
<223> Xaa equals stop translation
<400> 194
Met Gln Ala Gln Ile Ser Ser Pro Arg Trp Thr Ser Trp Phe Ser Leu
 1
                                     10
Thr Ala Val Thr Leu Ala Phe Pro Ser Leu Ile Pro Tyr Pro Ser Cys
                                 25
Gly Ile Pro Val Leu Thr Gln Asp Ala Lys Trp Pro Ser Asp Tyr Thr
                             40
Ser Pro Asp Ser Xaa
     50
<210> 195
<211> 186
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (114)
<223> Xaa equals any of the naturally occurring L-amino acids
<220>
<221> SITE
<222> (186)
<223> Xaa equals stop translation
<400> 195
Met Thr Leu Leu Asn Leu Leu Gln Thr Ile Phe Tyr Gly Val Thr
Cys Leu Asp Asp Val Leu Lys Arg Thr Lys Gly Gly Lys Asp Ile Lys
Phe Leu Thr Ala Phe Arg Asp Leu Leu Phe Thr Thr Leu Ala Phe Pro
Val Ser Thr Phe Val Phe Leu Ala Phe Trp Ile Leu Phe Leu Tyr Asn
Arg Asp Leu Ile Tyr Pro Lys Val Leu Asp Thr Val Ile Pro Val Trp
Leu Asn His Ala Met His Thr Phe Ile Phe Pro Ile Thr Leu Ala Glu
```

Val Val Leu Arg Pro His Ser Tyr Pro Ser Lys Lys Thr Gly Leu Thr
100 105 110

Leu Xaa Ala Ala Ala Ser Ile Ala Tyr Ile Ser Arg Ile Leu Trp Leu
115 120 125

Tyr Phe Glu Thr Gly Thr Trp Val Tyr Pro Val Phe Ala Lys Leu Ser 130 135 140

Leu Leu Gly Leu Ala Ala Phe Phe Ser Leu Ser Tyr Val Phe Ile Ala 145 150 155 160

Ser Ile Tyr Leu Leu Gly Glu Lys Leu Asn His Trp Lys Trp Gly Asp 165 170 175

Met Arg Gln Pro Arg Lys Lys Arg Lys Xaa 180 185

<210> 196

<211> 77

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (77)

<223> Xaa equals stop translation

<400> 196

Met Lys Asn Ala Thr Leu Leu Arg Met Val Leu Phe Val Ile Asn Leu 1 5 10 15

Gln Asn Leu Lys Ser Cys Pro Val Leu His Ile His Gln Asp Val Gln 20 25 30

Gln Gln Lys Arg Met Gly His Gly Gly Ser Ser Thr Arg Val Thr Val 35 40 45

Thr Ser Leu Ile Arg His Cys Thr Val Phe Gln Arg Pro Lys Asn Cys 50 55 60

Val Gln Asn Met Ile Thr Leu Gln Leu Ser Phe Pro Xaa 65 70 75

<210> 197

<211> 175

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (175)

<223> Xaa equals stop translation

<400> 197 Met Phe Val Pro Ser Cys Leu Cys Leu Arg Phe Val Val Thr Ser Leu Leu Leu Gln Met Thr His Ser Cys Gly Gly Phe Tyr Ile Cys Val Ile 25 Phe Glu Thr Ile Leu Ser Glu Phe Lys Thr Gln Ile Gly Arg Leu Tyr 40 Arg Lys Arg His Ile Gln Arg Lys Glu Ser Pro Lys Gly Arg Phe Val Met Leu Leu Pro Ser Ser Thr His Thr Ile Pro Phe Tyr Pro Asn Pro 70 Leu His Pro Arg Pro Phe Pro Ser Ser Arg Leu Pro Pro Gly Ile Ile Gly Gly Glu Tyr Asp Gln Arg Pro Thr Leu Pro Tyr Val Gly Asp Pro Ile Ser Ser Leu Ile Pro Gly Pro Gly Glu Thr Pro Ser Gln Phe Pro Pro Leu Arg Pro Arg Phe Asp Pro Val Gly Pro Leu Pro Gly Pro Asn 135 Pro Ile Leu Pro Gly Arg Gly Gly Pro Asn Asp Arg Phe Pro Phe Arg 155 Pro Ser Arg Gly Arg Pro Thr Asp Gly Arg Leu Ser Phe Met Xaa 170 <210> 198 <211> 51 <212> PRT <213> Homo sapiens <220> <221> SITE <222> (51) <223> Xaa equals stop translation Met Gly Leu Lys Arg Lys Gln Gly Phe Val Phe Leu Phe Leu Leu Leu

Lys Ser Thr Val Ala Ser Trp Leu Leu Ser Gly Val Gly Arg Ile Trp

Gly Leu Val His Phe Val Lys Val Asn His Val Cys Leu Asn Asn Arg

118

Gly Val Xaa 50

<210> 199

<211> 190

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (190)

<223> Xaa equals stop translation

<400> 199

Met Gly Pro Val Arg Leu Gly Ile Leu Leu Phe Leu Phe Leu Ala Val

1 5 10 15

His Glu Ala Trp Ala Gly Met Leu Lys Glu Glu Asp Asp Asp Thr Glu 20 25 30

Arg Leu Pro Ser Lys Cys Glu Val Cys Lys Leu Leu Ser Thr Glu Leu 35 40

Gln Ala Glu Leu Ser Arg Thr Gly Arg Ser Arg Glu Val Leu Glu Leu 50 60

Gly Gln Val Leu Asp Thr Gly Lys Arg Lys Arg His Val Pro Tyr Ser 65 70 75 80

Val Ser Glu Thr Arg Leu Glu Glu Ala Leu Glu Asn Leu Cys Glu Arg 85 90 95

Ile Leu Asp Tyr Ser Val His Ala Glu Arg Lys Gly Ser Leu Arg Tyr
100 105 110

Ala Lys Gly Gln Ser Gln Thr Met Ala Thr Leu Lys Gly Leu Val Gln 115 120 125

Lys Gly Val Lys Val Asp Leu Gly Ile Pro Leu Glu Leu Trp Asp Glu
130 135 140

Glu Glu Glu Glu Glu Glu Glu Glu Gly Gly Asp Lys Met Thr
165 170 175

Lys Thr Gly Ser His Pro Lys Leu Asp Arg Glu Asp Leu Xaa 180 185 190

<210> 200

<211> 80

Val Leu Tyr Val Cys Thr Leu Gly Thr Ala Cys Ile Trp Lys Ile Xaa

65 70 75 80

<210> 201 <211> 106 <212> PRT <213> Homo sapiens

<220>

<221> SITE

<222> (106)

<223> Xaa equals stop translation

<400> 201

Met Asn Leu Gly Val Ser Met Leu Arg Ile Leu Phe Leu Leu Asp Val 1 5 10 15

Gly Gly Ala Gln Val Leu Ala Thr Gly Lys Thr Pro Gly Ala Glu Ile 20 25 30

Asp Phe Lys Tyr Ala Leu Ile Gly Thr Ala Val Gly Val Ala Ile Ser 35 40 45

Ala Gly Phe Leu Ala Leu Lys Ile Cys Met Ile Arg Arg His Leu Phe 50 60

Asp Asp Asp Ser Ser Asp Leu Lys Ser Thr Pro Gly Gly Leu Ser Asp 65 70 75 80

Thr Ile Pro Leu Lys Lys Arg Ala Pro Arg Arg Asn His Asn Phe Ser

. 120

85 90 95

Lys Arg Asp Ala Gln Val Ile Glu Leu Xaa 100 105

<210> 202

<211> 80

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (80)

<223> Xaa equals stop translation

<400> 202

Met Ala Cys Leu Gly Gly Leu Leu Gly Ile Ile Gly Val Ile Cys Leu

1 5 10 15

Ile Ser Cys Leu Ser Pro Glu Met Asn Cys Asp Gly Gly His Ser Tyr
20 25 30

Val Arg Asn Tyr Leu Gln Lys Pro Thr Phe Ala Leu Gly Glu Leu Tyr 35 40 45

Pro Pro Leu Ile Asn Leu Trp Glu Ala Gly Lys Glu Lys Ser Thr Ser 50 55 60

Leu Lys Val Lys Ala Thr Val Ile Gly Leu Pro Thr Asn Met Ser Xaa 65 70 75 80

<210> 203

<211> 58

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (58)

<223> Xaa equals stop translation

<400> 203

Met Gly Leu Lys Leu Gln Lys Pro Gly Ser Leu Lys Thr Leu Ile

Ala Ile Ile Leu Val Met Tyr Ile Phe Met Thr Ile Ser Val Ile Ala 20 25 30

Gly Thr Gly Lys Phe Ser Gln Lys Leu Asp Leu His Leu Asn Met Asp 35 40

121

Ile Ser Pro Gly Arg Pro Ser Val His Xaa 50 55

<210> 204

<211> 161

<212> PRT

<213> Homo sapiens

<400> 204

Met Asp Phe Leu Asn Pro Asn Gly Ser Asp Cys Thr Leu Val Leu Phe 1 5 10 15 .

Tyr Thr Pro Trp Cys Arg Phe Ser Ala Ser Leu Ala Pro His Phe Asn 20 25 30

Ser Leu Pro Arg Ala Phe Pro Ala Leu His Phe Leu Ala Leu Asp Ala 35 40 45

Ser Gln His Ser Ser Leu Ser Thr Arg Phe Gly Thr Val Ala Val Pro 50 55 60

Asn Ile Leu Leu Phe Gln Gly Ala Lys Pro Met Ala Arg Phe Asn His 65 70 75 80

Thr Asp Arg Thr Leu Glu Thr Leu Lys Ile Phe Ile Phe Asn Gln Thr 85 90 95

Gly Ile Glu Ala Lys Lys Asn Val Val Val Thr Gln Ala Asp Gln Ile 100 105 110

Gly Pro Leu Pro Ser Thr Leu Ile Lys Ser Val Asp Trp Leu Leu Val 115 120 125

Phe Ser Leu Phe Phe Leu Ile Ser Phe Ile Met Tyr Ala Thr Ile Arg 130 135 140

Thr Glu Ser Ile Arg Trp Leu Ile Pro Gly Gln Glu Gln Glu His Val 145 150 155 160

Glu

<210> 205

<211> 137

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (10)

<223> Xaa equals any of the naturally occurring L-amino acids

122

<400> 205

Ile Pro Glu Asn Arg Arg Pro Ala Ser Xaa Cys Thr Trp Ser Met Trp 1 5 10 15

Thr Ser Arg Thr Thr Thr Arg Arg Pro Pro Trp Gly Arg Phe Ser Ser 20 25 30

Val Ser Ser Ala Ser Val Ser Ser Thr Arg Lys Thr Trp Arg Thr Arg 35 40 45

Ser Thr Ser Cys Cys Arg Ser Ser Arg Arg Arg Val Ala Ala Pro Phe 50 60

Cys Thr Pro Ser Ala Ser Thr Glu Pro Ser Ala Arg Met Glu Pro Pro 65 70 75 80

Leu Glu Leu Pro Val Val His Thr Phe Ser Phe Leu Thr Phe Val Phe 85 90 95

Thr Tyr Arg Cys Ser Ala Gly Asp Gly Ser Ile Thr Gln Ile Asn Cys 100 105 110

Ala Tyr Glu Met Gly Glu Glu Met Pro Lys Arg Gln Met Lys Ala Ile 115 120 125

Lys Phe Leu Leu Phe His Phe Tyr Leu

<210> 206

<211> 41

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (10)

<223> Xaa equals any of the naturally occurring L-amino acids.

<400> 206

Ile Pro Glu Asn Arg Arg Pro Ala Ser Xaa Cys Thr Trp Ser Met Trp 1 5 10 15

Thr Ser Arg Thr Thr Thr Arg Arg Pro Pro Trp Gly Arg Phe Ser Ser 20 25 30

Val Ser Ser Ala Ser Val Ser Ser Thr 35 40

<210> 207

<211> 43

<212> PRT

<213> Homo sapiens

123

<400> 207

Arg Lys Thr Trp Arg Thr Arg Ser Thr Ser Cys Cys Arg Ser Ser Arg

Arg Arg Val Ala Ala Pro Phe Cys Thr Pro Ser Ala Ser Thr Glu Pro

Ser Ala Arg Met Glu Pro Pro Leu Glu Leu Pro 40

<210> 208

<211> 53

<212> PRT

<213> Homo sapiens

<400> 208

Val Val His Thr Phe Ser Phe Leu Thr Phe Val Phe Thr Tyr Arg Cys

Ser Ala Gly Asp Gly Ser Ile Thr Gln Ile Asn Cys Ala Tyr Glu Met

Gly Glu Glu Met Pro Lys Arg Gln Met Lys Ala Ile Lys Phe Leu Leu

Phe His Phe Tyr Leu 50

<210> 209

<211> 223

<212> PRT

<213> Homo sapiens

<400> 209

His Pro Ser Ile Ile Ile Trp Ser Gly Asn Asn Glu Asn Glu Glu Ala 5 10

Leu Met Met Asn Trp Tyr His Ile Ser Phe Thr Asp Arg Pro Ile Tyr

Ile Lys Asp Tyr Val Thr Leu Tyr Val Lys Asn Ile Arg Glu Leu Val 35 40

Leu Ala Gly Asp Lys Ser Arg Pro Phe Ile Thr Ser Ser Pro Thr Asn

Gly Ala Glu Thr Val Ala Glu Ala Trp Val Ser Gln Asn Pro Asn Ser

Asn Tyr Phe Gly Asp Val His Phe Tyr Asp Tyr Ile Ser Asp Cys Trp

Asn Trp Lys Val Phe Pro Lys Ala Arg Phe Ala Ser Glu Tyr Gly Tyr

100 105 Gln Ser Trp Pro Ser Phe Ser Thr Leu Glu Lys Val Ser Ser Thr Glu 120 Asp Trp Ser Phe Asn Ser Lys Phe Ser Leu His Arg Gln His His Glu 135 Gly Gly Asn Lys Gln Met Leu Tyr Gln Ala Gly Leu His Phe Lys Leu 150 155 Pro Gln Ser Thr Asp Pro Leu Arg Thr Phe Lys Asp Thr Ile Tyr Leu 165 170 Thr Gln Val Met Gln Ala Gln Cys Val Lys Thr Glu Thr Glu Phe Tyr 180 185 Arg Arg Ser Arg Ser Glu Ile Val Asp Gln Gln Gly His Thr Met Gly 200 Ala Leu Tyr Trp Gln Leu Asn Asp Ile Trp Gln Ala Pro Ser Trp 215 <210> 210 <211> 160 <212> PRT <213> Homo sapiens <400> 210 Val Arg Val His Thr Trp Ser Ser Leu Glu Pro Val Cys Ser Arg Val Thr Glu Arg Phe Val Met Lys Gly Glu Ala Val Cys Leu Tyr Glu Glu Pro Val Ser Glu Leu Leu Arg Arg Cys Gly Asn Cys Thr Arg Glu Ser Cys Val Val Ser Phe Tyr Leu Ser Ala Asp His Glu Leu Leu Ser Pro Thr Asn Tyr His Phe Leu Ser Ser Pro Lys Glu Ala Val Gly Leu 70 Cys Lys Ala Gln Ile Thr Ala Ile Ile Ser Gln Gln Gly Asp Ile Phe Val Phe Asp Leu Glu Thr Ser Ala Val Ala Pro Phe Val Trp Leu Asp Val Gly Ser Ile Pro Gly Arg Phe Ser Asp Asn Gly Phe Leu Met Thr 120 Glu Lys Thr Arg Thr Ile Leu Phe Tyr Pro Trp Glu Pro Thr Ser Lys

125

130 135 140

Asn Glu Leu Glu Gln Ser Phe His Val Thr Ser Leu Thr Asp Ile Tyr 145 150 155 160

<210> 211

<211> 171

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (102)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 211

Pro Arg Leu Thr Pro Arg Met Lys Trp Pro Thr Ala Ala Leu Ala Ser 1 5 10 15

Arg Leu Leu Gly Trp Thr Val Leu Arg Pro Pro Tyr Pro Arg Val Pro 20 . 25 30

Ser Leu Pro Gln Val Thr Leu His Pro Thr Asp Gly Leu Met Ala Val 35 40 45

Leu Tyr Thr Gly Gly Glu Gly Arg Thr Leu Gly Glu Gln His Phe Phe 50 60

His Glu Thr Phe Val Thr Arg Trp Leu Leu Gly Pro Val Pro Val Arg
65 70 75 80

Phe Gly Ala Cys Ser Pro Leu Ser Phe Leu Ala Pro Arg Arg Gly Gln 85 90 95

Gly Ala Pro Ala Gly Xaa Phe Cys Ala Cys Pro Arg Pro Ala Ser Arg 100 105 110

Gln Leu Cys Pro Trp Pro Ala Leu Pro Gly Thr Pro Tyr Ser Asn Ser 115 120 125

Ala Pro Leu Cys Thr Gly Met Gly His Ser Asn Thr Pro Gln Gly Pro 130 135 140

Pro Ser Pro Gln Tyr Ala Leu Ser Pro Thr Glu Pro Thr Ser Leu Ser 145 150 155 160

Gly Asn Ser His Leu Pro Ala Ile Leu Val Leu 165 170

<210> 212

```
<211> 41
<212> PRT
```

<213> Homo sapiens

<400> 212

Pro Arg Leu Thr Pro Arg Met Lys Trp Pro Thr Ala Ala Leu Ala Ser 1 5 10 15

Arg Leu Leu Gly Trp Thr Val Leu Arg Pro Pro Tyr Pro Arg Val Pro 20 25 30

Ser Leu Pro Gln Val Thr Leu His Pro 35 40

<210> 213

<211> 41

<212> PRT

<213> Homo sapiens

<400> 213

Thr Asp Gly Leu Met Ala Val Leu Tyr Thr Gly Gly Glu Gly Arg Thr 1 5 10 15

Leu Gly Glu Gln His Phe Phe His Glu Thr Phe Val Thr Arg Trp Leu 20 25 30

Leu Gly Pro Val Pro Val Arg Phe Gly 35

<210> 214

<211> 42

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (20)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 214

Ala Cys Ser Pro Leu Ser Phe Leu Ala Pro Arg Arg Gly Gln Gly Ala 1 5 10 15

Pro Ala Gly Xaa Phe Cys Ala Cys Pro Arg Pro Ala Ser Arg Gln Leu 20 25 30

Cys Pro Trp Pro Ala Leu Pro Gly Thr Pro 35

<210> 215

<211> 47

<212> PRT

127

<213> Homo sapiens

<400> 215

Tyr Ser Asn Ser Ala Pro Leu Cys Thr Gly Met Gly His Ser Asn Thr 1 5 10 15

Pro Gln Gly Pro Pro Ser Pro Gln Tyr Ala Leu Ser Pro Thr Glu Pro
20 25 30

Thr Ser Leu Ser Gly Asn Ser His Leu Pro Ala Ile Leu Val Leu 35 40 45

<210> 216

<211> 27

<212> PRT

<213> Homo sapiens

<400> 216

His Leu Leu Glu Val Thr Pro Cys Arg Leu Pro Val Pro Glu Phe Pro 1 5 10 15

Gly Arg Thr Pro Arg Gly Ser Arg Thr Pro Asp
20 25

<210> 217

<211> 239

<212> PRT

<213> Homo sapiens

<400> 217

Met Ile Pro Gly Ser Asp Ser Gln Thr Ala Leu Asn Phe Gly Ser Thr

1 5 10 15

Leu Met Lys Lys Lys Ser Asp Pro Glu Gly Pro Ala Leu Leu Phe Pro 20 . 25 30

Glu Ser Glu Leu Ser Ile Arg Ile Gly Arg Ala Gly Leu Leu Ser Asp 35 40 45

Lys Ser Glu Asn Gly Glu Ala Tyr Gln Arg Lys Lys Ala Ala Ala Thr 50 55 60

Gly Leu Pro Glu Gly Pro Ala Val Pro Val Pro Ser Arg Gly Asn Leu 65 70 75 80

Ala Gln Pro Gly Gly Ser Ser Trp Arg Arg Ile Ala Leu Leu Ile Leu 85 90 95

Ala Ile Thr Ile His Asn Val Pro Glu Gly Leu Ala Val Gly Val Gly 100 105 110

Phe Gly Ala Ile Glu Lys Thr Ala Ser Ala Thr Phe Glu Ser Ala Arg 115 120 125

128

Asn Leu Ala Ile Gly Ile Gly Ile Gln Asn Phe Pro Glu Gly Leu Ala 130 135 140

Val Ser Leu Pro Leu Arg Gly Ala Gly Phe Ser Thr Trp Arg Ala Phe 145 150 155 160

Trp Tyr Gly Gln Leu Ser Gly Met Val Glu Pro Leu Ala Gly Val Phe 165 · 170 175

Gly Ala Phe Ala Val Val Leu Ala Glu Pro Ile Leu Pro Tyr Ala Leu 180 185 190

Ala Phe Ala Ala Gly Ala Met Val Tyr Val Val Met Asp Asp Ile Ile 195 200 205

Pro Glu Ala Gln Ile Ser Gly Asn Gly Lys Leu Ala Ser Trp Ala Ser 210 215 220

Ile Leu Gly Phe Val Val Met Met Ser Leu Asp Val Gly Leu Gly 225 230 235

<210> 218

<211> 43

<212> PRT

<213> Homo sapiens

<400> 218

Met Ile Pro Gly Ser Asp Ser Gln Thr Ala Leu Asn Phe Gly Ser Thr 1 5 10 15

Leu Met Lys Lys Lys Ser Asp Pro Glu Gly Pro Ala Leu Leu Phe Pro 20 25 30

Glu Ser Glu Leu Ser Ile Arg Ile Gly Arg Ala
35 40

<210> 219

<211> 41

<212> PRT

<213> Homo sapiens

<400> 219

Gly Leu Leu Ser Asp Lys Ser Glu Asn Gly Glu Ala Tyr Gln Arg Lys
1 5 10 15

Lys Ala Ala Ala Thr Gly Leu Pro Glu Gly Pro Ala Val Pro Val Pro 20 25 30

Ser Arg Gly Asn Leu Ala Gln Pro Gly 35

129

```
<210> 220
<211> 44
<212> PRT
<213> Homo sapiens
<400> 220
Gly Ser Ser Trp Art
1
```

Gly Ser Ser Trp Arg Arg Ile Ala Leu Leu Ile Leu Ala Ile Thr Ile 1 5 10 15

His Asn Val Pro Glu Gly Leu Ala Val Gly Val Gly Phe Gly Ala Ile $20 \hspace{1cm} 25 \hspace{1cm} 30$

Glu Lys Thr Ala Ser Ala Thr Phe Glu Ser Ala Arg

<210> 221 <211> 43 <212> PRT <213> Homo sapiens

<400> 221

Asn Leu Ala Ile Gly Ile Gly Ile Gln Asn Phe Pro Glu Gly Leu Ala 1 5 10 15

Val Ser Leu Pro Leu Arg Gly Ala Gly Phe Ser Thr Trp Arg Ala Phe 20 25 30

Trp Tyr Gly Gln Leu Ser Gly Met Val Glu Pro 35

<210> 222 <211> 43 <212> PRT <213> Homo sapiens

<400> 222
Leu Ala Gly Val Phe Gly Ala Phe Ala Val Val Leu Ala Glu Pro Ile
1 5 10 15

Leu Pro Tyr Ala Leu Ala Phe Ala Ala Gly Ala Met Val Tyr Val Val 20 25 30

Met Asp Asp Ile Ile Pro Glu Ala Gln Ile Ser 35

<210> 223 <211> 25 <212> PRT <213> Homo sapiens

 $<\!400\!>$ 223 Gly Asn Gly Lys Leu Ala Ser Trp Ala Ser Ile Leu Gly Phe Val Val

130

1 10 15 Met Met Ser Leu Asp Val Gly Leu Gly 20 <210> 224 <211> 11 <212> PRT <213> Homo sapiens <400> 224 Thr Arg Pro Ile Thr Tyr Val Leu Leu Ala Gly 5 <210> 225 <211> 35 <212> PRT <213> Homo sapiens <400> 225 Gly Thr Ser Leu Thr Ala Pro Leu Leu Glu Phe Leu Leu Ala Leu Tyr Phe Leu Phe Ala Asp Ala Met Gln Leu Asn Asp Lys Trp Gln Gly Leu Cys Trp Pro <210> 226 <211> 30 <212> PRT <213> Homo sapiens <400> 226 Leu Ala Asn Phe Glx Cys Ser Asp Cys Ala Gln Thr Val Leu Phe Val Leu Glx Phe Glx Ile Leu Val Phe Thr Tyr Glu Ile Pro Phe 25 20 <210> 227 <211> 75 <212> PRT <213> Homo sapiens <400> 227 Gln Ala Trp His Glu Val Gly Gly Val Arg Arg Cys Trp Phe Val

Leu Gly Glu Arg Arg Ala Gly Ser Leu Leu Ser Ala Ser Tyr Gly Thr

131

20 25 30

Phe Ala Met Pro Gly Met Val Leu Phe Gly Arg Arg Trp Ala Ile Ala 35 40 45

Ser Asp Asp Leu Val Phe Pro Gly Phe Phe Glu Leu Val Val Arg Val 50 60

Leu Trp Trp Ile Gly Ile Leu Thr Leu Tyr Leu 65 70 75

<210> 228

<211> 125

<212> PRT

<213> Homo sapiens

<400> 228

Pro Gly Met Val Leu Phe Gly Arg Arg Trp Ala Ile Ala Ser Asp Asp 1 5 10 15

Leu Val Phe Pro Gly Phe Phe Glu Leu Val Val Arg Val Leu Trp Trp
20 25 30

Ile Gly Ile Leu Thr Leu Tyr Leu Met His Arg Gly Lys Leu Asp Cys
35 40 45

Ala Gly Gly Ala Leu Leu Ser Ser Tyr Leu Ile Val Leu Met Ile Leu 50 55 60

Leu Ala Val Val Ile Cys Thr Val Ser Ala Ile Met Cys Val Ser Met 65 70 75 80

Arg Gly Thr Ile Cys Asn Pro Gly Pro Arg Lys Ser Met Ser Lys Leu $85 \hspace{1cm} 90 \hspace{1cm} 95$

Leu Tyr Ile Arg Leu Ala Leu Phe Phe Pro Glu Met Val Trp Ala Ser 100 105 110

Leu Gly Ala Ala Trp Val Ala Asp Gly Val Gln Cys Asp 115 120 125

<210> 229

<211> 18

<212> PRT

<213> Homo sapiens

<400> 229

His Glu Arg Asn Cys Phe Pro Met Trp Leu Asn His Ser Ala Phe Pro 1 5 10 15

Pro Val

<210> 230

<211> 132

<212> PRT

<213> Homo sapiens

<400> 230

Gly Trp Thr Arg Glu Asn Asp His Arg Ala Leu Ser Lys Ala Gly Ile 1 5 10 15

Gly Ser Ala Glu Ile Gln Pro Ser Asn Leu Arg Val Gly Ser Ala Lys 20 25 30

Asp Leu Gly Lys Pro Trp Ala Gly Lys Leu Leu Leu Leu Ser Ser Cys
35 40 45

Leu Leu Phe Phe Ser Leu Gly Val Leu Tyr Arg Gly Gln Met Leu Ala 50 60

Pro Pro Leu Gln Glu Asp Trp Lys Gly Gly Val Lys Asp Ser Asp Leu 65 70 75 80

Ile Asp Asp Ser Ser Ala Ser Pro Ile Pro Pro Ser Tyr Leu Glu Tyr 85 90 95

Lys Ala Ala Leu Tyr Pro Phe Ser Glu His Lys Ser Val Arg Asn Ala 100 105 110

Thr Asp Ser Leu Thr Phe Phe Leu Val Thr Asp His Phe Leu Asp Asn 115 120 125

Gln Asp Ser Gln 130

<210> 231

<211> 45

<212> PRT

<213> Homo sapiens

<400> 231

Gly Trp Thr Arg Glu Asn Asp His Arg Ala Leu Ser Lys Ala Gly Ile 1 5 10 15

Gly Ser Ala Glu Ile Gln Pro Ser Asn Leu Arg Val Gly Ser Ala Lys 20 25 30

Asp Leu Gly Lys Pro Trp Ala Gly Lys Leu Leu Leu Leu 35 40 45

<210> 232

<211> 46

<212> PRT

<213> Homo sapiens

<400> 232 Ser Ser Cys Leu Leu Phe Phe Ser Leu Gly Val Leu Tyr Arg Gly Gln Met Leu Ala Pro Pro Leu Gln Glu Asp Trp Lys Gly Gly Val Lys Asp Ser Asp Leu Ile Asp Asp Ser Ser Ala Ser Pro Ile Pro Pro 35 40 <210> 233 <213> Homo sapiens <400> 233 Ser Tyr Leu Glu Tyr Lys Ala Ala Leu Tyr Pro Phe Ser Glu His Lys Ser Val Arg Asn Ala Thr Asp Ser Leu Thr Phe Phe Leu Val Thr Asp 25 His Phe Leu Asp Asn Gln Asp Ser Gln <210> 234 <211> 11 <212> PRT <213> Homo sapiens <400> 234 Leu Lys Phe His Gln Glu Ser Leu Ser Gly Asp 1 5 . 10 <210> 235 <211> 25 <212> PRT <213> Homo sapiens <400> 235 Glu Ala Lys Ser Arg Pro Val Thr Gln Ala Gly Val Gln Trp His Asp 10 5 Leu Gly Ser Leu Gln Pro Leu Pro Pro <210> 236 <211> 25 <212> PRT <213> Homo sapiens

<400> 236

Glu Ala Lys Ser Arg Pro Val Thr Gln Ala Gly Val Gln Trp His Asp

Leu Gly Ser Leu Gln Pro Leu Pro Pro 20

<210> 237

<211> 137

<212> PRT

<213> Homo sapiens

<400> 237

Ala Leu Val Leu Val Cys Arg Gln Arg Tyr Cys Arg Pro Arg Asp Leu

Leu Gln Arg Tyr Asp Ser Lys Pro Ile Val Asp Leu Ile Gly Ala Met

Glu Thr Gln Ser Glu Pro Ser Glu Leu Glu Leu Asp Asp Val Val Ile

Thr Asn Pro His Ile Glu Ala Ile Leu Glu Asn Glu Asp Trp Ile Glu

Asp Ala Ser Gly Leu Met Ser His Cys Ile Ala Ile Leu Lys Ile Cys

His Thr Leu Thr Glu Lys Leu Val Ala Met Thr Met Gly Ser Gly Ala

Lys Met Lys Thr Ser Ala Ser Val Ser Asp Ile Ile Val Val Ala Lys 105

Arg Ile Ser Pro Arg Val Asp Asp Val Val Lys Ser Met Tyr Pro Pro 115 120 125

Leu Asp Pro Lys Leu Leu Asp Ala Arg 130 135

<210> 238

<211> 319

<212> PRT

<213> Homo sapiens

<400> 238

Asp Val Glu Ser Arg Gly Pro Ser Ala Arg Cys Leu Pro Val Val Pro

Gly Ser Leu Leu Pro Gly Leu Glu Pro Ala Thr Lys Leu Met Pro Gly 25

Gly	Leu	Ala 35	Pro	Gly	His	Gly	Ala 40	Pro	Val	Arg	Glu	Leu 45	Leu	Leu	Pro
Leu	Leu 50	Ser	Gln	Pro	Thr	Leu 55	Gly	Ser	Leu	Trp	Asp 60	Ser	Leu	Arg	His
Суs 65	Ser	Leu	Leu	Cys	Asn 70	Pro	Leu	Ser	Cys	Val 75	Pro	Ala	Leu	Glu	Ala 80
Pro	Pro	Ser	Leu	Val 85	Ser	Leu	Gly	Cys	Ser 90	Gly	Gly	Cys	Pro	Arg 95	Leu
Ser	Leu	Ala	Gly 100	Ser	Ala	Ser	Pro	Phe 105	Pro	Phe	Leu	Thr	Ala 110	Leu	Leu
Ser	Leu	Leu 115	Asn	Thr	Leu	Ala	Gln 120	Ile	His	Lys	Gly	Leu 125	Cys	Gly	Glr
Leu	Ala 130	Ala	Ile	Leu	Ala	Ala 135	Pro	Gly	Leu	Gln	Asn 140	Tyr	Phe	Leu	Gln
Cys 145	Val	Ala	Pro	Gly	Ala 150	Ala	Pro	His	Leu	Thr 155	Pro	Phe	Ser	Ala	Trp 160
Ala	Leu	Arg	His	Glu 165	Tyr	His	Leu	Gln	Tyr 170	Leu	Ala	Leu	Ala	Leu 175	Ala
Gln	Lys	Ala	Ala 180	Ala	Leu	Gln	Pro	Leu 185	Pro	Ala	Thr	His	Ala 190	Ala	Leu
Tyr	His	Gly 195	Met	Ala	Leu	Ala	Leu 200	Leu	Ser	Arg	Leu	Leu 205	Pro	Gly	Ser
Glu	Tyr 210	Leu	Thr	His	Glu	Leu 215	Leu	Leu	Ser	Cys	Val 220	Phe	Arg	Leu	Glu
Phe 225	Leu	Pro	Glu	Arg	Thr 230	Ser	Gly	Gly	Pro	Glu 235	Ala	Ala	Asp	Phe	Ser 240
Asp	Gln	Leu	Ser	Leu 245	Gly	Ser	Ser	Arg	Val 250	Pro	Arg	Cys	Gly	Gln 255	Gly
Thr	Leu	Leu	Ala 260	Gln	Ala	Cys	Gln	Asp 265	Leu	Pro	Ser	Ile	Arg 270		Cys
Tyr	Leu	Thr 275	His	Cys	Ser	Pro	Ala 280	Arg	Ala	Ser	Leu	Leu 285	Ala	Ser	Gln
Ala	Leu 290	His	Arg	Gly	Glu	Leu 295	Gln	Arg	Val	Pro	Thr 300	Leu	Leu	Leu	Pro
Met 305	Pro	Thr	Glu	Pro	Leu 310	Leu	Pro	Thr	Asp	Trp 315	Pro	Phe	Leu	His	

```
<210> 239
<211> 44
<212> PRT
<213> Homo sapiens
<400> 239
Asp Val Glu Ser Arg Gly Pro Ser Ala Arg Cys Leu Pro Val Val Pro
Gly Ser Leu Leu Pro Gly Leu Glu Pro Ala Thr Lys Leu Met Pro Gly
                               25
Gly Leu Ala Pro Gly His Gly Ala Pro Val Arg Glu
                            40
<210> 240
<211> 45
<212> PRT
<213> Homo sapiens
Leu Leu Pro Leu Leu Ser Gln Pro Thr Leu Gly Ser Leu Trp Asp
                 5
Ser Leu Arg His Cys Ser Leu Leu Cys Asn Pro Leu Ser Cys Val Pro
Ala Leu Glu Ala Pro Pro Ser Leu Val Ser Leu Gly Cys
<210> 241
<211> 45
<212> PRT
<213> Homo sapiens
<400> 241
Ser Gly Gly Cys Pro Arg Leu Ser Leu Ala Gly Ser Ala Ser Pro Phe
Pro Phe Leu Thr Ala Leu Leu Ser Leu Leu Asn Thr Leu Ala Gln Ile
                                25
His Lys Gly Leu Cys Gly Gln Leu Ala Ala Ile Leu Ala
        35 40
<210> 242
<211> 44
```

\Z11> 44

<212> PRT

<213> Homo sapiens

<400> 242

Ala Pro Gly Leu Gln Asn Tyr Phe Leu Gln Cys Val Ala Pro Gly Ala

1 10 15 Ala Pro His Leu Thr Pro Phe Ser Ala Trp Ala Leu Arg His Glu Tyr 25 20 His Leu Gln Tyr Leu Ala Leu Ala Leu Ala Gln Lys 35 40 <210> 243 <211> 44 <212> PRT <213> Homo sapiens <400> 243 Ala Ala Leu Gln Pro Leu Pro Ala Thr His Ala Ala Leu Tyr His 5 10 Gly Met Ala Leu Ala Leu Leu Ser Arg Leu Leu Pro Gly Ser Glu Tyr 20 25 Leu Thr His Glu Leu Leu Leu Ser Cys Val Phe Arg 40 <210> 244 <211> 44 <212> PRT <213> Homo sapiens <400> 244 Leu Glu Phe Leu Pro Glu Arg Thr Ser Gly Gly Pro Glu Ala Ala Asp Phe Ser Asp Gln Leu Ser Leu Gly Ser Ser Arg Val Pro Arg Cys Gly 25 Gln Gly Thr Leu Leu Ala Gln Ala Cys Gln Asp Leu 40 <210> 245 <211> 53 <212> PRT <213> Homo sapiens <400> 245 Pro Ser Ile Arg Asn Cys Tyr Leu Thr His Cys Ser Pro Ala Arg Ala 5 10 Ser Leu Leu Ala Ser Gln Ala Leu His Arg Gly Glu Leu Gln Arg Val 25

Pro Thr Leu Leu Pro Met Pro Thr Glu Pro Leu Pro Thr Asp 40

35

```
Trp Pro Phe Leu His
    50
<210> 246
<211> 25
<212> PRT
<213> Homo sapiens
<400> 246
Val Gly Ser Val Leu Gly Ala Phe Leu Thr Phe Pro Gly Leu Arg Leu
                                     10
Ala Gln Thr His Arg Asp Ala Leu Thr
             20
<210> 247
<211> 65
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (21)
<223> Xaa equals any of the naturally occurring L-amino acids
<220>
<221> SITE
<222> (37)
<223> Xaa equals any of the naturally occurring L-amino acids
<220>
<221> SITE
<222> (48)
<223> Xaa equals any of the naturally occurring L-amino acids
<220>
<221> SITE
<222> (57)
<223> Xaa equals any of the naturally occurring L-amino acids
<400> 247
Leu Glu Cys Thr Asp Thr Ile Met Val His Cys Ser Leu Lys Leu Leu
                                     10
Ser Pro Ser Asp Xaa Ser His Ser Ala Ser Gln Val Ala Lys Thr Arg
Gly Val His His Xaa Thr Gln Leu Ile Phe Lys Val Phe Phe Val Xaa
         35
                             40
Met Gly Ser His Ser Thr Lys Tyr Xaa Ser Ile Arg Pro Gly Leu Leu
     50
                         55
                                             60
```

Pro 65 <210> 248 <211> 14 <212> PRT <213> Homo sapiens <400> 248 Glu Ser Ser Phe Val Pro Pro Ala Ala His Ser Ser Leu Cys 5 10 <210> 249 <211> 172 <212> PRT <213> Homo sapiens <220> <221> SITE <222> (72) <223> Xaa equals any of the naturally occurring L-amino acids <400> 249 Leu Leu Pro Gly Gln Gln Glu Ala Thr Gln Cys Val Glu Ala Gly Ala 10 Gly Glu Gly Ala Leu Thr Pro Met Cys Pro Cys Arg Gln Glu Gln Phe Val Asp Leu Tyr Lys Glu Phe Glu Pro Ser Leu Val Asn Ser Thr Val Tyr Ile Met Ala Met Ala Ile Gln Met Ala Pro Phe Ala Ile Asn Tyr Lys Val Arg Pro Gly Pro Cys Xaa Asn Ile His Cys Leu Pro Thr Gln Pro His Pro Met Lys Pro Ser Val Pro His Pro His Arg Ala Arg Pro Ser Trp Arg Ala Cys Pro Arg Thr Ser Pro Trp Cys Gly Val Trp Gln 105 Phe His Ser Trp Pro Ser Leu Ala Cys Ser Ser Ala Pro Arg Pro Thr 115 120 Ser Thr Ala Ser Leu Ala Ser Trp Thr Ser Leu Trp Ser Ser Ser Trp 135 Ser Leu Pro Arg Ser Cys Ser Trp Thr Ser Ala Trp Arg Ser Trp Pro 155 145 150

```
Thr Ala Ser Cys Ser Ser Ser Trp Gly Pro Arg Ser 165 170
```

<210> 250

<211> 45

<212> PRT

<213> Homo sapiens

<400> 250

Leu Leu Pro Gly Gln Gln Glu Ala Thr Gln Cys Val Glu Ala Gly Ala

1 10 15

Gly Glu Gly Ala Leu Thr Pro Met Cys Pro Cys Arg Gln Glu Gln Phe 20 25 30

Val Asp Leu Tyr Lys Glu Phe Glu Pro Ser Leu Val Asn 35 40 45

<210> 251

<211> 44

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (27)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 251

Ser Thr Val Tyr Ile Met Ala Met Ala Ile Gln Met Ala Pro Phe Ala 1 5 10 15

Ile Asn Tyr Lys Val Arg Pro Gly Pro Cys Xaa Asn Ile His Cys Leu 20 25 30

Pro Thr Gln Pro His Pro Met Lys Pro Ser Val Pro 35

<210> 252

<211> 42

<212> PRT

<213> Homo sapiens

<400> 252

His Pro His Arg Ala Arg Pro Ser Trp Arg Ala Cys Pro Arg Thr Ser 1 10 15

Pro Trp Cys Gly Val Trp Gln Phe His Ser Trp Pro Ser Leu Ala Cys
20 25 30

Ser Ser Ala Pro Arg Pro Thr Ser Thr Ala

35 40

<210> 253

<211> 41

<212> PRT

<213> Homo sapiens

<400> 253

Ser Leu Ala Ser Trp Thr Ser Leu Trp Ser Ser Ser Trp Ser Leu Pro 1 10 15

Arg Ser Cys Ser Trp Thr Ser Ala Trp Arg Ser Trp Pro Thr Ala Ser 20 25 30

Cys Ser Ser Ser Trp Gly Pro Arg Ser 35 40

<210> 254

<211> 48

<212> PRT

<213> Homo sapiens

<400> 254

Thr Arg Asn Ile Leu Ser Phe Ile Lys Cys Val Ile His Asn Phe Trp 1 5 10 15

Ile Pro Lys Glu Ser Asn Glu Ile Thr Ile Ile Ile Asn Pro Tyr Arg

Glu Thr Val Cys Phe Ser Val Glu Pro Val Lys Lys Ile Phe Asn Tyr 35 40 45

<210> 255

<211> 27

<212> PRT

<213> Homo sapiens

<400> 255

Leu Val Val Leu Phe Ala Ser Ser Asn Ser Arg Tyr Leu Lys Tyr Phe 1 5 10 15

Phe Leu Val Pro Leu Ile Leu Gly Ser Ala Trp
20 25

<210> 256

<211> 20

<212> PRT

WO 99/47540

142

```
<400> 256
His Glu Trp Lys Cys Lys Gln Lys Tyr Ser Glu Gly Ser Gly Asn Thr
                                    10
Arg Ile Gly Asn
            20
<210> 257
<211> 20
<212> PRT
<213> Homo sapiens
<400> 257
Leu Leu Pro Leu Cys Phe Leu Gly Pro Arg Gln Val Leu Glu Glu Phe
Pro Ser Ile Val
<210> 258
<211> 12
<212> PRT
<213> Homo sapiens
<400> 258
Pro Thr Arg Pro Ser Lys His Gln Glu Ala Gly Ser
<210> 259
<211> 42
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (39)
<223> Xaa equals any of the naturally occurring L-amino acids
<400> 259
Gly Gln Gly Pro Ala Gly Arg Trp Val Arg Arg Leu Pro Cys Ser Arg
Arg Ala Gly Glu Arg Gly Pro His Trp Gly Val Trp Ala Gly Pro
            20
                                25
Gln Met Ser Cys Gly Leu Xaa Phe Gly Pro
         35
                             40
<210> 260
<211> 193
```

<212> PRT <213> Homo sapiens

<400> 260

Trp Arg Thr Gln Gly Pro Met Val Leu Leu Trp Val Val Thr Cys Pro
1 5 10 15

Ala Thr Met Leu Thr Glu Pro Gln Asn Pro His Leu Ile Gly Phe Val 20 25 30

Ala Tyr Ser Gly Pro Ser His Thr Thr Gln Pro His Lys Tyr Trp Leu 35 40 45

Leu Leu Asp Gly Gln Ala Asp Pro Ala Ala Ala Glu Gly Pro Val Lys 50 60

Arg Lys Ala Ala Ser Val Val Trp Trp Pro Gln Ala Leu Arg His Leu 65 70 75 80

Ser Leu Leu Val His Cys Trp Glu Glu Ser Tyr Glu Met Asn Ile Gly 85 90 95

Cys Gln Ser Leu Trp Ala Gly Gly Leu Ala Ser Ser Gly Asn Gly Trp
100 105 110

Asp Leu Gly Val Ala Phe Arg Arg Asp Thr Cys Met Ser Ser Ser Ser 115 120 125

Leu His Trp Lys Glu Phe Lys Tyr Ala Pro Gly Ser Leu His Tyr Phe 130 135 140

Ala Leu Ser Phe Val Leu Ile Leu Thr Glu Ile Cys Leu Val Ser Ser 145 150 155 160

Gly Met Gly Phe Pro Gln Glu Gly Lys His Phe Ser Val Leu Gly Ser 165 170 175

Pro Asp Cys Ser Leu Trp Gly Arg Asp Glu His Val Pro Arg Glu Phe 180 185 190

Ala

<210> 261

<211> 42

<212> PRT

<213> Homo sapiens

<400> 261

Trp Arg Thr Gln Gly Pro Met Val Leu Leu Trp Val Val Thr Cys Pro 1 5 10 15

Ala Thr Met Leu Thr Glu Pro Gln Asn Pro His Leu Ile Gly Phe Val 20 25 30

Ala Tyr Ser Gly Pro Ser His Thr Thr Gln 35 40

<210> 262

<211> 41

<212> PRT

<213> Homo sapiens

<400> 262

Pro His Lys Tyr Trp Leu Leu Leu Asp Gly Gln Ala Asp Pro Ala Ala 1 5 10 15

Ala Glu Gly Pro Val Lys Arg Lys Ala Ala Ser Val Val Trp Pro 20 25 30

Gln Ala Leu Arg His Leu Ser Leu Leu
35 40

<210> 263

<211> 41

<212> PRT

<213> Homo sapiens

<400> 263

Val His Cys Trp Glu Glu Ser Tyr Glu Met Asn Ile Gly Cys Gln Ser

Leu Trp Ala Gly Gly Leu Ala Ser Ser Gly Asn Gly Trp Asp Leu Gly 20 . 25 30

Val Ala Phe Arg Arg Asp Thr Cys Met

<210> 264

<211> 44

<212> PRT

<213> Homo sapiens

<400> 264

Ser Ser Ser Leu His Trp Lys Glu Phe Lys Tyr Ala Pro Gly Ser 1 5 10 15

Leu His Tyr Phe Ala Leu Ser Phe Val Leu Ile Leu Thr Glu Ile Cys
20 25 30

Leu Val Ser Ser Gly Met Gly Phe Pro Gln Glu Gly

<210> 265

<211> 25

<212> PRT

<213> Homo sapiens

<400> 265

Lys His Phe Ser Val Leu Gly Ser Pro Asp Cys Ser Leu Trp Gly Arg
1 5 10 15

Asp Glu His Val Pro Arg Glu Phe Ala 20 25

<210> 266

<211> 31

<212> PRT

<213> Homo sapiens

<400> 266

Ile Ala Gln Gly Thr Val Pro Leu Thr Lys Arg Gly Val Gln Ser Ser 1 5 10 15

Gly Pro Asp Tyr Pro Glu Gly Thr Leu Thr Pro Leu Pro Arg Gly 20 25 30

<210> 267

<211> 31

<212> PRT

<213> Homo sapiens

<400> 267

Ile Ala Gln Gly Thr Val Pro Leu Thr Lys Arg Gly Val Gln Ser Ser 1 5 10 15

Gly Pro Asp Tyr Pro Glu Gly Thr Leu Thr Pro Leu Pro Arg Gly 20 25 30

<210> 268

<211> 28

<212> PRT

<213> Homo sapiens

<400> 268

Asp Cys Leu Tyr Leu Ala Leu Ser Phe Pro Trp His Cys His Cys His 1 10 15

His His Pro Pro Ser Gly Ser Leu Leu Tyr Pro Phe 20 25

<210> 269

<211> 101

<212> PRT

146

<400> 269

Ala Ser Leu Pro Pro Ser Arg Ser Arg Pro Leu Ala Asn Met Ala Leu

1 5 10 15

Val Pro Cys Gln Val Leu Arg Met Ala Ile Leu Leu Ser Tyr Cys Ser 20 25 30

Ile Leu Cys Asn Tyr Lys Ala Ile Glu Met Pro Ser His Gln Thr Tyr $35 \hspace{1.5cm} 40 \hspace{1.5cm} 45$

Gly Gly Ser Trp Lys Phe Leu Thr Phe Ile Asp Leu Val Ile Gln Ala 50 55 60

Val Phe Phe Gly Ile Cys Val Leu Thr Asp Leu Ser Ser Leu Leu Thr 65 70 75 80

Arg Gly Ser Gly Asn Gln Glu Gln Glu Arg Gln Leu Lys Lys Leu Ile 85 90 95

Ser Leu Arg Asp Trp 100

<210> 270

<211> 16

<212> PRT

<213> Homo sapiens

<400> 270

Met Ser Arg Ser Ser Arg Ile Ser Gly Leu Ser Cys Pro Trp Leu Leu 1 5 10

<210> 271

<211> 45

<212> PRT

<213> Homo sapiens

<400> 271

Asp His Trp Pro Ala Gly Phe Leu Pro Pro Ala Pro Gly Leu Lys Phe 1 5 10 15

Pro Val Ala Leu Glu Val Phe Arg Lys Val Leu Pro Ala Val Cys Pro 20 25 30

Thr Asp Cys Ser Gly Ser Ala Gly Lys Glu Arg Asn Ser 35 40 45

<210> 272

<211> 47

<212> PRT

147

```
<213> Homo sapiens
<400> 272
Glu Glu Ile Ala Thr S
```

Glu Glu Ile Ala Thr Ser Ile Glu Pro Ile Arg Asp Phe Leu Ala Ile 1 5 10 15

Val Phe Phe Ala Ser Ile Gly Leu His Val Phe Pro Thr Phe Val Ala 20 25 30

Tyr Glu Leu Thr Val Leu Val Phe Leu Thr Leu Ser Val Val Val 35 40 45

<210> 273 <211> 7

<211> / <212> PRT

<213> Homo sapiens

<400> 273

Tyr Cys Asn Leu Gln Cys Arg

<210> 274

<211> 44

<212> PRT

<213> Homo sapiens

<400> 274

Ser Ala Leu Ile Gly Asn Pro Lys Gly Cys Phe Gly Cys Phe Ser Pro 1 5 10 15

Val Val Leu Arg Glu Trp Ser Val Glu Ser Trp Lys Ser Leu Arg Pro 20 25 30

Phe Gln Ala Ile Cys Lys Leu Lys Thr Asn Phe Arg 35

<210> 275

<211> 8

<212> PRT

<213> Homo sapiens

<400> 275

His Glu Ala Ala Leu Arg Gly Pro 1 5

<210> 276

<211> 26

<212> PRT

<213> Homo sapiens

<400> 276

Ser Asn Ala Ala Gly Asn Val Val Arg Ala Phe Leu Tyr Ile Asn His 1 5 10 15

Leu Lys Leu Gly Cys Lys Val Gly Leu Ala 20 25

<210> 277

<211> 25

<212> PRT

<213> Homo sapiens

<400> 277

Asn Trp Ala Val Leu Asn Met Leu Leu Ser Lys Gly Lys Ile Thr Ile
1 5 10 15

Phe Leu Gly Pro Leu Glu Cys Gly Ser 20 25

<210> 278

<211> 49

<212> PRT

<213> Homo sapiens

<400> 278

Pro Ser His Gln Thr Arg Lys Gly Lys Ser Ala Lys Leu Leu Asp Arg

1 5 10 15

Pro Pro Glu Ala Leu Arg Met Lys Ile Ile Thr Thr Thr Leu Leu Leu 20 25 30

Ala Cys His Leu Gln Leu Glu Val Gly Val Val Gly Gly Glu Val
35 40 45

Asp

<210> 279

<211> 51

<212> PRT

<213> Homo sapiens

<400> 279

Phe Gln Ala Ser Ser Ala Asn Asn Gln Gln Asn Trp Gly Ser Gln Pro 1 5 10

Ile Ala Gln Gln Pro Leu Gln Gln Gly Gly Asp Tyr Ser Gly Asn Tyr
20 25 30

Gly Tyr Asn Asn Asp Asn Gln Glu Phe Tyr Gln Asp Thr Tyr Gly Gln 35 40 45

Gln Trp Lys

<210> 280 <211> 264 <212> PRT <213> Homo sapiens <220> <221> SITE <222> (2) <223> Xaa equals any of the naturally occurring L-amino acids <220> <221> SITE . <222> (6) <223> Xaa equals any of the naturally occurring L-amino acids <220> <221> SITE <222> (14) <223> Xaa equals any of the naturally occurring L-amino acids <400> 280 Trp Xaa Pro Leu Leu Xaa Thr Ser Gly Ser Pro Gly Leu Xaa Gly Phe 5 10 15 Gly Thr Arg Met Asn Gly Lys Glu Ile Glu Gly Glu Glu Ile Glu Ile 25 Val Leu Ala Lys Pro Pro Asp Lys Lys Arg Lys Glu Arg Gln Ala Ala 35 40 45 Arg Gln Ala Ser Arg Ser Thr Ala Tyr Glu Asp Tyr Tyr Tyr His Pro Pro Pro Arg Met Pro Pro Pro Ile Arg Gly Arg Gly Arg Gly Gly Arg Gly Gly Tyr Gly Tyr Pro Pro Asp Tyr Tyr Gly Tyr Glu Asp Tyr Tyr Asp Asp Tyr Tyr Gly Tyr Asp Tyr His Asp Tyr Arg Gly Gly Tyr 105 Glu Asp Pro Tyr Tyr Gly Tyr Asp Asp Gly Tyr Ala Val Arg Gly Arg Gly Gly Arg Gly Gly Arg Gly Ala Pro Pro Pro Pro Arg Gly Arg Gly Ala Pro Pro Pro Arg Gly Arg Ala Gly Tyr Ser Gln Arg Gly Ala 145 155 Pro Leu Gly Pro Pro Arg Gly Ser Arg Gly Gly Arg Gly Pro Ala

150

165 170 175 Gln Gln Gln Arg Gly Arg Gly Ser Arg Gly Ser Arg Gly Asn Arg Gly 185 Gly Asn Val Gly Gly Lys Arg Lys Ala Asp Gly Tyr Asn Gln Pro Asp 200 Ser Lys Arg Arg Gln Pro Thr Thr Asn Arg Thr Gly Val Pro Asn Pro 215 Ser Leu Ser Ser Arg Phe Ser Lys Val Val Thr Ile Leu Val Thr Met 230 235 Val Thr Ile Met Thr Thr Arg Asn Phe Ile Arg Ile Leu Met Gly Asn 245 250 Ser Gly Ser Arg Gln Val Arg Ala 260 <210> 281 <211> 27 <212> PRT <213> Homo sapiens <400> 281 Arg Met Asn Gly Lys Glu Ile Glu Gly Glu Glu Ile Glu Ile Val Leu Ala Lys Pro Pro Asp Lys Lys Arg Lys Glu Arg 20 <210> 282 <211> 25 <212> PRT <213> Homo sapiens <400> 282 Tyr Tyr His Pro Pro Pro Arg Met Pro Pro Pro Ile Arg Gly Arg Gly Arg Gly Gly Arg Gly Gly Tyr Gly 20 <210> 283 <211> 26 <212> PRT <213> Homo sapiens Asp Tyr Arg Gly Gly Tyr Glu Asp Pro Tyr Tyr Gly Tyr Asp Asp Gly

10

151

```
Tyr Ala Val Arg Gly Arg Gly Gly Gly Arg
    20 25
<210> 284
<211> 28
<212> PRT
<213> Homo sapiens
<400> 284
Pro Pro Pro Arg Gly Arg Ala Gly Tyr Ser Gln Arg Gly Ala Pro Leu
Gly Pro Pro Arg Gly Ser Arg Gly Gly Arg Gly Gly
           20
                               25
<210> 285
<211> 35
<212> PRT
<213> Homo sapiens
<400> 285
Ala Asp Gly Tyr Asn Gln Pro Asp Ser Lys Arg Arg Gln Pro Thr Thr
                                 10
1
Asn Arg Thr Gly Val Pro Asn Pro Ser Leu Ser Ser Arg Phe Ser Lys
            20
                               25
Val Val Thr
        35
<210> 286
<211> 19
<212> PRT
<213> Homo sapiens
<400> 286
Leu Gln Ile Pro Pro Ser Ser Gln Ser Leu Gly Leu Lys Asn Ala Asp
1 5
Ser Ser Ile
<210> 287
<211> 129
<212> PRT
<213> Homo sapiens
Gly Gly Pro Pro Glu Ser Ala Pro Trp Leu Pro Ala Val Leu Arg Ala
```

152

Pro Val Leu Thr Ser Arg Cys Ala Ser Ser Asp Ser Glu Gly Pro Val

Trp Phe Cys Gln Pro Gly Ser Gly Pro Ser Ser Thr Glu Met Ser Cys
35 40 45

His Cys Ile Leu Gly Pro Gly Ser Ser Cys Leu Cys Val Leu Arg Gly 50 55 60

Ser Met Trp Thr Pro Ser Val Pro Gly Trp Pro Gln Pro Ala Lys Glu 65 70 75 80

Thr Gly Ala Ser Ser Cys Ser Val Phe Ser Ala Asn Asn Gly Ser Cys
85 90 95

Pro Leu Pro Leu His Asn His Gln Arg Gln Ala Ser Leu Asp Thr Gly
100 105 110

Leu Ser Leu Glu His Val Pro Gly Glu Ser Tyr Phe Tyr Ser Pro Val 115 120 125

Gly

<210> 288

<211> 34

<212> PRT

<213> Homo sapiens

<400> 288

Ser Ser Asp Ser Glu Gly Pro Val Trp Phe Cys Gln Pro Gly Ser Gly 1 5 10 15

Pro Ser Ser Thr Glu Met Ser Cys His Cys Ile Leu Gly Pro Gly Ser 20 25 30

Ser Cys

<210> 289

<211> 28

<212> PRT

<213> Homo sapiens

<400> 289

Trp Thr Pro Ser Val Pro Gly Trp Pro Gln Pro Ala Lys Glu Thr Gly
1 5 10 15

Ala Ser Ser Cys Ser Val Phe Ser Ala Asn Asn Gly
20 25

```
<210> 290
<211> 21
<212> PRT
<213> Homo sapiens
Gln Arg Gln Ala Ser Leu Asp Thr Gly Leu Ser Leu Glu His Val Pro
Gly Glu Ser Tyr Phe
            20
<210> 291
<211> 29
<212> PRT
<213> Homo sapiens
<400> 291
Ser Ser Ser Leu Val Leu Thr Ile Arg Ser Gln Thr Leu Phe Leu Ala
                                    10
Ser Phe Ile His Ser Thr Ser Ile Phe Cys Ala Leu Asn
            20
                                25
<210> 292
<211> 12
<212> PRT
<213> Homo sapiens
<400> 292
Cys Cys Cys Arg Leu Gly Leu Ser Gly Pro Lys Cys
1 5
<210> 293
<211> 22
<212> PRT
<213> Homo sapiens
<400> 293
Arg Ala Phe Trp Gly Leu Gly Ala Leu Gln Leu Leu Asp Leu Ser Ala
Asn Gln Leu Glu Ala Leu
            20
<210> 294
<211> 34
<212> PRT
<213> Homo sapiens
<400> 294
```

His Ala Ser Gly Arg Arg Thr Gly Ser Ala Asp Asp Gly Leu Gln Gly
1 5 10 15

Arg Thr Gly Ser Gly Pro Pro Thr Ala Gly Ala Gly Gly Gly Ala 20 25 30

Ala Pro

<210> 295

<211> 205

<212> PRT

<213> Homo sapiens

<400> 295

Val Ser Ala Ala Ala Gly Ala Arg Leu Ala Pro Arg Ala Pro Gly Ala 1 5 10 15

Pro Ala Gly Cys Arg Pro Met Arg Gly Cys Ala Ala Arg Ala Ala Ala 20 25 30

Arg Lys Ser Leu Val Pro Val Leu Pro Ala Gly Trp Arg Ser Gly Pro 35 40 45

Ala Ala Ala Arg Pro Gly Pro Arg Arg Leu Ala His Ala Pro Ser 50 55 60

Ala Ala Arg Ser Arg Ala Gly Pro Gly Ala Val Ala Arg Pro Leu Pro 65 70 75 80

Arg Arg His Leu Ala Ala His Gly Arg Gly Cys Gly Pro Ala Ala 85 90 95

Ala Arg Ala Gly Ala Gly Ser Gly Pro Gly Ala Arg Arg Ala Ala Arg 100 105 110

Val Pro Thr Ala Gly Arg Pro Pro Gly Thr His Val His Thr Ser Gly 115 120 125

Gln Ser Gly Ala Pro Arg Asp Pro Glu Gly Glu Ala Leu Ala Asp Thr 130 135 140

Trp Ala Gln Thr Gly Gln Gly Asp Ser Ser Ser Asn Ser Ser Ser Ser 145 150 155 160

Gly Arg Gly Arg Asp Gln Glu Gly Pro Arg Met Gly Ala Ala Pro Pro 165 170 175

Pro Pro Ala Pro Ala Val Gly Gly Pro Leu Pro Val Arg Pro Trp Ser 180 185 190

Pro Ser Ser Ala Glu Pro Val Leu Arg Pro Asp Ala Trp
195 200 205

<210> 296

<211> 368 <212> PRT <213> Homo sapiens <400> 296 Thr Arg Pro Ala Ala Glu Arg Ala Pro Arg Thr Thr Gly Ser Arg Asp 10 Ala Gln Ala Ala Gly Leu Pro Pro Arg Val Pro Gly Ala Gly Gly Leu Pro Pro Cys Gly Ala Leu Pro Gly Arg Gly Leu Gly Arg Cys Cys Cys Cys Cys Cys Cys Arg Leu Gly Leu Ser Gly Pro Lys Cys Arg Pro Gly Pro Arg Pro Arg Gly Pro Trp Ala Pro Arg Thr Ala Pro Arg Cys Ala Arg Ala Cys Arg Glu Ala Cys Gln Leu Ser Ala Leu Ser Leu Pro Ala Val Pro Pro Gly Leu Ser Leu Arg Leu Arg Ala Leu Leu Leu Asp 105 His Asn Arg Val Arg Ala Leu Pro Pro Gly Ala Phe Ala Gly Ala Gly 115 120 Ala Leu Gln Arg Leu Asp Leu Arg Glu Asn Gly Leu His Ser Val His 135 Val Arg Ala Phe Trp Gly Leu Gly Ala Leu Gln Leu Leu Asp Leu Ser 145 150 155 Ala Asn Gln Leu Glu Ala Leu Ala Pro Gly Thr Phe Ala Pro Leu Arg 165 170 Ala Leu Arg Asn Leu Ser Leu Ala Gly Asn Arg Leu Ala Arg Leu Glu 180 185 Pro Ala Ala Leu Gly Ala Leu Pro Leu Leu Arg Ser Leu Ser Leu Gln Asp Asn Glu Leu Ala Ala Leu Ala Pro Gly Leu Leu Gly Arg Leu Pro 215 Ala Leu Asp Ala Leu His Leu Arg Gly Asn Pro Trp Gly Cys Gly Cys Ala Leu Arg Pro Leu Cys Ala Trp Leu Arg Arg His Pro Leu Pro Ala 250

156

Ser Glu Ala Glu Thr Val Leu Cys Val Trp Pro Gly Arg Leu Thr Leu 260 265 270

Ser Pro Leu Thr Ala Phe Ser Asp Ala Ala Phe Ser His Cys Ala Gln 275 280 285

Pro Leu Ala Leu Arg Asp Leu Ala Arg Gly Leu His Ala Arg Ala Gly 290 295 300

Leu Leu Pro Arg Gln Pro Gly Phe Leu Pro Gly Ala Gly Leu Trp Ala 305 310 315 320

His Arg Leu Pro Cys Ala Pro Pro Pro Pro Pro His Arg Arg Pro Pro 325 330 335

Pro Ala Glu Thr Val Gln Thr Arg Thr Pro Ile Pro Thr Pro Thr Ala 340 345 350

Val Pro Arg Pro Arg Thr Arg Gly Ala Pro Ser Ala Ala Ala Gln Ala 355 360 365

<210> 297

<211> 47

<212> PRT

<213> Homo sapiens

<400> 297

Gly Cys Arg Pro Met Arg Gly Cys Ala Ala Arg Ala Ala Ala Arg Lys

1 5 10 15

Ser Leu Val Pro Val Leu Pro Ala Gly Trp Arg Ser Gly Pro Ala Ala . 20 25 30

Ala Ala Arg Pro Gly Pro Arg Arg Leu Ala His Ala Pro Ser Ala 35 40 45

<210> 298

<211> 30

<212> PRT

<213> Homo sapiens

<400> 298

Pro Gly Ala Val Ala Arg Pro Leu Pro Arg Arg His Leu Ala Ala Ala 1 5 10 15

His Gly Arg Gly Cys Gly Pro Ala Ala Ala Arg Ala Gly Ala 20 25 30

<210> 299

157

```
<211> 24
<212> PRT
<213> Homo sapiens
<400> 299
Ser Gly Gln Ser Gly Ala Pro Arg Asp Pro Glu Gly Glu Ala Leu Ala
Asp Thr Trp Ala Gln Thr Gly Gln
             20
<210> 300
<211> 23
<212> PRT
<213> Homo sapiens
<400> 300
Pro Pro Ala Pro Ala Val Gly Gly Pro Leu Pro Val Arg Pro Trp Ser
                                    10
Pro Ser Ser Ala Glu Pro Val
             20
<210> 301
<211> 26
<212> PRT
<213> Homo sapiens
<400> 301
Ala Pro Arg Thr Thr Gly Ser Arg Asp Ala Gln Ala Ala Gly Leu Pro
Pro Arg Val Pro Gly Ala Gly Gly Leu Pro
             20
<210> 302
<211> 22
<212> PRT
<213> Homo sapiens
<400> 302
Gly Pro Arg Pro Arg Gly Pro Trp Ala Pro Arg Thr Ala Pro Arg Cys
Ala Arg Ala Cys Arg Glu
             20
<210> 303
<211> 31
<212> PRT
```

<400> 303

Ala Val Pro Pro Gly Leu Ser Leu Arg Leu Arg Ala Leu Leu Leu Asp 1 5 10 15

His Asn Arg Val Arg Ala Leu Pro Pro Gly Ala Phe Ala Gly Ala 20 25 30

<210> 304

<211> 24

<212> PRT

<213> Homo sapiens

<400> 304

Leu Gly Ala Leu Gln Leu Leu Asp Leu Ser Ala Asn Gln Leu Glu Ala 1 5 10

Leu Ala Pro Gly Thr Phe Ala Pro 20

<210> 305

<211> 36

<212> PRT

<213> Homo sapiens

<400> 305

Pro Pro Gly Ala Phe Ala Gly Ala Gly Ala Leu Gln Arg Leu Asp Leu 1 5 10

Arg Glu Asn Gly Leu His Ser Val His Val Arg Ala Phe Trp Gly Leu 20 25 30

Gly Ala Leu Gln 35

<210> 306

<211> 28

<212> PRT

<213> Homo sapiens

<400> 306

Arg Asn Leu Ser Leu Ala Gly Asn Arg Leu Ala Arg Leu Glu Pro Ala 1 5 10 15

Ala Leu Gly Ala Leu Pro Leu Leu Arg Ser Leu Ser 20 25

<210> 307

<211> 26

<212> PRT

<400> 307

Leu Pro Ala Leu Asp Ala Leu His Leu Arg Gly Asn Pro Trp Gly Cys

Gly Cys Ala Leu Arg Pro Leu Cys Ala Trp 20

<210> 308

<211> 34

<212> PRT

<213> Homo sapiens

<400> 308

Thr Val Leu Cys Val Trp Pro Gly Arg Leu Thr Leu Ser Pro Leu Thr

Ala Phe Ser Asp Ala Ala Phe Ser His Cys Ala Gln Pro Leu Ala Leu 20

Arg Asp

<210> 309

<211> 24

<212> PRT

<213> Homo sapiens

<400> 309

Leu His Ala Arg Ala Gly Leu Leu Pro Arg Gln Pro Gly Phe Leu Pro 10

Gly Ala Gly Leu Trp Ala His Arg 20

<210> 310

<211> 24

<212> PRT

<213> Homo sapiens

<400> 310

Thr Val Gln Thr Arg Thr Pro Ile Pro Thr Pro Thr Ala Val Pro Arg 5

Pro Arg Thr Arg Gly Ala Pro Ser 20

<210> 311

<211> 59

<212> PRT

160 <400> 311 His Ala Ser Gly Arg Pro Asp Arg Ser Ser Ala Pro Ile Gly Asn Ser Gly Leu Pro Cys Pro Asp Leu Glu Pro Leu Gly Gly Leu Gln Ser Lys Cys Arg Leu Cys Ala Pro Thr Glu Ala Arg Gly Leu Trp Ser Arg Ser Leu Cys Ser Asp Arg Cys Asp Thr Trp Arg Ser <210> 312 <211> 29 <212> PRT <213> Homo sapiens <400> 312 Gly Leu Pro Cys Pro Asp Leu Glu Pro Leu Gly Gly Leu Gln Ser Lys 10 1 5 Cys Arg Leu Cys Ala Pro Thr Glu Ala Arg Gly Leu Trp . 25 20 <210> 313 <211> 16 <212> PRT <213> Homo sapiens

Gln Glu Trp Glu Ser Glu Leu Gly Glu Arg Arg Lys Pro Leu Gln Ala 1 5 10 15

<210> 314 <211> 46 <212> PRT <213> Homo sapiens <400> 314

Cys Gln Ser Ser Asn Leu Ile Phe Phe Gln Phe Val Asn Ile Leu Phe 1 5 10 15

Asn Leu Met Met Asp Ile Leu Val Asp Phe Ser Ile Thr Lys Met Pro 20 25 30

Ile Asn Ser Ile Phe Ser Leu Tyr Phe Cys Tyr Glu Ile Ile 35 40 45

<210> 315

<211> 134

<212> PRT

<213> Homo sapiens

<400> 315

Gly Pro Val Trp Leu Phe Cys Phe Leu Thr Leu Cys Arg Lys Pro Ser

Gln Leu Phe Ser Gln Glu Asn Ser Cys Met Asp Val Ala Gly Gly Val

Thr Thr Cys Leu Pro Pro Trp Phe Ser Arg Gly Ala Pro Ala Gln Met

Ser Gln Trp Pro Pro Ser Ser Asp His Gly Ala Val Arg Ala Gly Arg

Asp Ser Arg Val Gly Pro Val Gln Pro Ser His Leu Thr Cys Glu Gly

Gly Lys Glu Glu Arg Glu Lys Asn Lys Lys Ala Glu Val Asn Pro Pro 90

Thr Gly Met Gly Leu Ala Asn Arg Ile Pro Arg Asp Asp Ile Thr Leu

Lys Leu Arg Asn Gln Gly Lys Leu Arg Thr Lys Glu Asn Arg Thr Gln 120

Ser Ala Lys Arg His Pro 130

<210> 316

<211> 42

<212> PRT

<213> Homo sapiens

<400> 316

Val Ala Cys Lys Pro Glu Asn Arg Thr Lys Thr His Phe Ala Ser Ser 10

Pro Ala Cys Asp Gly His Ala Leu Gly Gly Gln Val Gly Phe Ala Ile

Cys Phe Leu Ser Cys Leu Phe Pro Pro Met

<210> 317

<211> 40

<212> PRT

```
<213> Homo sapiens
```

<400> 317

Ser His Pro Met Pro Asn Thr Pro Gln Lys Gln Leu Leu Phe Ser Glu

Asp Asn Glu Leu Leu Val Ser Leu Arg Thr Gly Arg Lys Pro Thr Leu

Gln Ala Ala Leu Arg Val Thr Gly 35 40

<210> 318

<211> 59

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (26)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 318

Glu Gly Asp Pro Arg Gly Arg Pro Arg Pro Arg Pro Leu Gly Pro Pro

Pro Gln Leu Thr Leu Pro Thr Ala Leu Xaa Asp Ile Leu Arg Gln Val

Arg Ala Pro Gly Leu Arg Leu Ser Arg Ala Leu Glu Val Gly Arg Lys

Gly Ser Pro Ile Phe Lys Ile Gln Ile Tyr Leu 55

<210> 319

<211> 250

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (145)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 319

Ala His Arg Leu Gln Ile Arg Leu Leu Thr Trp Asp Val Lys Asp Thr

Leu Leu Arg Leu Arg His Pro Leu Gly Glu Ala Tyr Ala Thr Lys Ala

Arg Ala His Gly Leu Glu Val Glu Pro Ser Ala Leu Glu Gln Gly Phe

		35					40					45			
Arg	Gln 50	Ala	Tyr	Arg	Ala	Gln 55	Ser	His	Ser	Phe	Pro 60	Asn	Tyr	Gly	Leu
Ser 65	His	Gly	Leu	Thr	Ser 70	Arg	Gln	Trp	Trp	Leu 75	Asp	Val	Val	Leu	Gln 80
Thr	Phe	His	Leu	Ala 85	Gly	Val	Gln	Asp	Ala 90	Gln	Ala	Val	Ala	Pro 95	Ile
Ala	Glu	Gln	Leu 100	Tyr	Lys	Asp	Phe	Ser 105	His	Pro	Cys	Thr	Trp 110	Gln	Val
Leu	Asp	Gly 115	Ala	Glu	Asp	Thr	Leu 120	Arg	Glu	Cys	Arg	Thr 125	Arg	Gly	Leu
Arg	Leu 130	Ala	Val	Ile	Ser	Asn 135	Phe	Asp	Arg	Arg	Leu 140	Glu	Gly	Ile	Leu
Xaa 145	Gly	Leu	Gly	Leu	Arg 150	Glu	His	Phe	Asp	Phe 155	Val	Leu	Thr	Ser	Glu 160
Ala	Ala	Gly	Trp	Pro 165	Lys	Pro	Asp	Pro	Arg 170	Ile	Phe	Gln	Glu	Ala 175	Leu
Arg	Leu	Ala	His 180	Met	Glu	Pro	Val	Val 185	Ala	Ala	His	Val	Gly 190	Asp	Asn
Tyr	Leu	Cys 195	Asp	Tyr	Gln	Gly	Pro 200	Arg	Ala	Val	Gly	Met 205	His	Ser	Phe
Leu	Val 210		Gly	Pro	Gln	Ala 215	Leu	Asp	Pro	Val	Val 220	Arg	Asp	Ser	Val
Pro 225	-	Glu	His	Ile	Leu 230	Pro	Ser	Leu	Ala	His 235		Leu	Pro	Ala	Leu 240
Asp	Cys	Leu	Glu	Gly 245		Thr	Pro	Gly	Leu 250						
<210> 320 <211> 27 <212> PRT <213> Homo sapiens															
			. Leu	Thr 5		Asp	Val	. Lys	Asp 10		Leu	Leu	Arg	Leu 15	

His Pro Leu Gly Glu Ala Tyr Ala Thr Lys Ala 20 $\,\,$ 25

```
<210> 321
<211> 24
<212> PRT
<213> Homo sapiens
<400> 321
Leu Glu Gln Gly Phe Arg Gln Ala Tyr Arg Ala Gln Ser His Ser Phe
Pro Asn Tyr Gly Leu Ser His Gly
           20
<210> 322
<211> 26
<212> PRT
<213> Homo sapiens
<400> 322
His Leu Ala Gly Val Gln Asp Ala Gln Ala Val Ala Pro Ile Ala Glu
Gln Leu Tyr Lys Asp Phe Ser His Pro Cys
            20
<210> 323
<211> 23
<212> PRT
<213> Homo sapiens
Val Leu Asp Gly Ala Glu Asp Thr Leu Arg Glu Cys Arg Thr Arg Gly
                                    10
Leu Arg Leu Ala Val Ile Ser
           20
<210> 324
<211> 26
<212> PRT
<213> Homo sapiens
<400> 324
Arg Glu His Phe Asp Phe Val Leu Thr Ser Glu Ala Ala Gly Trp Pro
Lys Pro Asp Pro Arg Ile Phe Gln Glu Ala
            20
```

<210> 325 <211> 28

<212> PRT

165

```
<213> Homo sapiens
```

<400> 325

Glu Pro Val Val Ala Ala His Val Gly Asp Asn Tyr Leu Cys Asp Tyr 1 5 10 15

Gln Gly Pro Arg Ala Val Gly Met His Ser Phe Leu 20 25

<210> 326

<211> 23

<212> PRT

<213> Homo sapiens

<400> 326

Val Val Arg Asp Ser Val Pro Lys Glu His Ile Leu Pro Ser Leu Ala 1 5 10 15

His Leu Leu Pro Ala Leu Asp 20

<210> 327

<211> 22

<212> PRT

<213> Homo sapiens

<400> 327

Ile Arg Lys Leu Gly Pro Gly Leu Ala Pro Cys Ser Cys Arg Ser Gly
1 5 10 15

Gln Val Phe Pro Arg Val 20

<210> 328

<211> 241

<212> PRT

<213> Homo sapiens

<400> 328

Lys Pro Leu Arg Met Ala Arg Pro Gly Gly Pro Glu His Asn Glu Tyr

1 10 15

Ala Leu Val Ser Ala Trp His Ser Ser Gly Ser Tyr Leu Asp Ser Glu 20 25 30

Gly Leu Arg His Gln Asp Asp Phe Asp Val Ser Leu Leu Val Cys His $35 \hspace{1cm} 40 \hspace{1cm} 45$

Cys Ala Ala Pro Phe Glu Glu Glu Glu Glu Ala Glu Arg His Val Leu $50 \hspace{1.5cm} 55 \hspace{1.5cm} 60$

Arg Leu Gln Phe Phe Val Val Leu Thr Ser Gln Arg Glu Leu Phe Pro

65					70					75					80
Arg	Leu	Thr	Ala	Asp 85	Met	Arg	Arg	Phe	Arg 90	Lys	Pro	Pro	Arg	Leu 95	Pro
Pro	Glu	Pro	Glu 100	Ala	Pro	Gly	Ser	Ser 105	Ala	Gly	Ser	Pro	Gly 110	Glu	Ala
Ser	Gly	Leu 115	Ile	Leu	Ala	Pro	Gly 120	Pro	Ala	Pro	Leu	Phe 125	Pro	Pro	Leu
Ala	Ala 130	Glu	Val	Gly	Met	Ala 135	Arg	Ala	Arg	Leu	Ala 140	Gln	Leu	Val	Arg
Leu 145	Ala	Gly	Gly	His	Cys 150	Arg	Arg	Asp	Thr	Leu 155	Trp	Lys	Arg	Leu	Phe 160
Leu	Leu	Glu	Pro	Pro 165	Gly	Pro	Asp	Arg	Leu 170	Arg	Leu	Gly	Gly	Arg 175	Leu
Ala	Leu	Ala	Glu 180	Leu	Glu	Glu	Leu	Leu 185	Glu	Ala	Val	His	Ala 190	Lys	Ser
Ile	Gly	Asp 195	Ile	Asp	Pro	Gln	Leu 200	Asp	Cys	Phe	Leu	Ser 205	Met	Thr	Val
Ser	Trp 210	Tyr	Gln	Ser	Leu	Ile 215	Lys	Val	Leu	Leu	Ser 220	Arg	Phe	Pro	Arg
Ala 225	Val	Ala	Ile	Ser	Lys 230	Ala	Gln	Thr	Trp	Glu 235	Leu	Ser	Thr	Trp	Leu 240
Arg															
<210> 329 <211> 30 <212> PRT <213> Homo sapiens															
)> 32 Arg		Thr	Leu 5	Glu	Leu	Pro	Thr	Pro 10	Leu	Ile	Ala	Ala	His 15	Gln
Leu	Tyr	Asn	Tyr 20	Val	Ala	Asp	His	Ala 25	Ser	Ser	Tyr	His	Met 30		
<210> 330 <211> 37 <212> PRT <213> Homo sapiens															

<400> 330

167

Ser His Cys Glu Trp Pro Gly Gln Gly Ala Gln Asn Thr Thr Ser Met

1 5 10 15

Pro Trp Cys Arg His Gly Thr Val Leu Ala Pro Thr Trp Thr Leu Arg
20 25 30

Asp Phe Asp Thr Arg

<210> 331

<211> 91

<212> PRT

<213> Homo sapiens

<400> 331

Pro Leu Thr Thr Val Ser His Leu Cys Pro Leu Ser Leu Arg Val Phe 1 5 10 15

Thr Ser His Leu Asp Ile Thr Ala Gly His Ser His Arg Asp Asp Thr 20 25 30

Trp Val Pro Ile Pro Ala Leu Pro Leu Lys His Leu Arg Pro Pro Ser 35 40 45

Ser Pro Phe Ala Leu Gly Pro Trp Val Ser His Pro Leu Met Arg Trp 50 55 60

Val Gln Lys Leu Ser His Leu His Ser Asn Pro Gly Thr Gly Phe Ser 65 70 75 80

Met Gly Gly Lys Ser Ala Glu Lys Leu Lys Cys 85 90

<210> 332

<211> 179

<212> PRT

<213> Homo sapiens

<400> 332

Ser Thr Ala Ala Arg Gly Ala Pro Gly Pro Gly Arg Ala Gly Gly Thr 1 5 10

Pro Arg Ser Ser Pro Cys Gln Ile His Trp Gly His Arg Pro Pro Ala 20 25 30

Gly Leu Leu Pro Ile His Asp Gly Leu Leu Val Pro Glu Pro Asp Gln
35 40 45

Ser Ser Pro Lys Pro Leu Pro Gln Ser Cys Arg His Phe Gln Ser Pro 50 60

Asp Leu Gly Thr Gln Tyr Leu Val Ala Leu Asn Gln Lys Phe Thr Asp 65 70 75 80

168

Cys Ser Ala Leu Val Phe Trp Thr Pro Leu Arg Lys Asp Val Ser Glu 85 90 95

Val Val Phe Arg Glu Ala Leu Pro Val Gln Pro Gln Asp Thr Arg Ser 100 105 110

Pro Pro Ala Gln Leu Val Ser Thr Tyr His His Leu Glu Ser Val Ile 115 120 125

Asn Thr Ala Cys Phe Thr Leu Leu Asp Pro Pro Pro Leu Lys Gly Val 130 140

Asp Trp Thr Thr Glu Cys His Cys Ser Leu Asn His Gly Pro Thr Arg 145 150 155 160

Leu Pro Ala Arg Gly Arg Thr Asp Gln Pro Phe Trp Ala Pro Gly Gln
165 170 175

Ala Arg His

<210> 333

<211> 56

<212> PRT

<213> Homo sapiens

<400> 333

His Gln Arg Leu Cys Asn Tyr Val Leu Arg Val Cys Cys Pro Ser Leu 1 5 10

Ala Ala Gly Thr Ala Leu Pro Lys His Pro Gln Pro Leu Thr His Pro 20 25 30

Gly Leu Gln Arg Val Arg Ser Thr Pro Arg Thr Pro Trp Ala Leu Leu 35 40 45

Gly Tyr Ser Phe Arg Pro Pro Trp
50 55

<210> 334

<211> 28

<212> PRT

<213> Homo sapiens

<400> 334

Pro Gly Gly Pro Glu His Asn Glu Tyr Ala Leu Val Ser Ala Trp His

1 10 15

Ser Ser Gly Ser Tyr Leu Asp Ser Glu Gly Leu Arg 20 25

169

```
<210> 335
<211> 25
<212> PRT
<213> Homo sapiens
<400> 335
Asp Val Ser Leu Leu Val Cys His Cys Ala Ala Pro Phe Glu Glu Gln
Gly Glu Ala Glu Arg His Val Leu Arg
            20
<210> 336
<211> 28
<212> PRT
<213> Homo sapiens
<400> 336
Arg Leu Thr Ala Asp Met Arg Arg Phe Arg Lys Pro Pro Arg Leu Pro
Pro Glu Pro Glu Ala Pro Gly Ser Ser Ala Gly Ser
                  25
<210> 337
<211> 25
<212> PRT
<213> Homo sapiens
<400> 337
Gly Glu Ala Ser Gly Leu Ile Leu Ala Pro Gly Pro Ala Pro Leu Phe
                 5
1
Pro Pro Leu Ala Ala Glu Val Gly Met
            20
<210> 338
<211> 23
<212> PRT
<213> Homo sapiens
<400> 338
Thr Leu Trp Lys Arg Leu Phe Leu Leu Glu Pro Pro Gly Pro Asp Arg
                                  10
Leu Arg Leu Gly Gly Arg Leu
```

<210> 339 <211> 28 <212> PRT

PCT/US99/05804

<213> Homo sapiens

<400> 339

Leu Ala Glu Leu Glu Glu Leu Leu Glu Ala Val His Ala Lys Ser Ile
1 5 10 15

Gly Asp Ile Asp Pro Gln Leu Asp Cys Phe Leu Ser

<210> 340

<211> 197

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (97)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 340

Phe Gln Leu Tyr Phe Asn Pro Glu Leu Ile Phe Lys His Phe Gln Ile

1 5 10 15

Trp Arg Leu Ile Thr Asn Phe Leu Phe Phe Gly Pro Val Gly Phe Asn 20 25 30

Phe Leu Phe Asn Met Ile Phe Leu Tyr Arg Tyr Cys Arg Met Leu Glu 35 40

Glu Gly Ser Phe Arg Gly Arg Thr Ala Asp Phe Val Phe Met Phe Leu 50 60

Phe Gly Gly Phe Leu Met Thr Leu Phe Gly Leu Phe Val Ser Leu Val 65 70 75 80

Phe Leu Gly Gln Ala Phe Thr Ile Met Leu Val Tyr Val Trp Ser Arg 85 90 95

Xaa Asn Pro Tyr Val Arg Met Asn Phe Phe Gly Leu Leu Asn Phe Gln 100 105 110

Ala Pro Phe Leu Pro Trp Val Leu Met Gly Phe Ser Leu Leu Leu Gly 115 120 125

Asn Ser Ile Ile Val Asp Leu Leu Gly Ile Ala Val Gly His Ile Tyr 130 135 140

Phe Phe Leu Glu Asp Val Phe Pro Asn Gln Pro Gly Gly Ile Arg Ile 145 150 150 155

Leu Lys Thr Pro Ser Ile Leu Lys Ala Ile Phe Asp Thr Pro Asp Glu 165 170 175

Asp Pro Asn Tyr Asn Pro Leu Pro Glu Glu Arg Pro Gly Gly Phe Ala

171

180 185 190

Trp Gly Glu Gly Gln 195

<210> 341

<211> 108

<212> PRT

<213> Homo sapiens

<400> 341

Gly Val Gly Gln Ala Thr Val Gly Lys Met Ala Tyr Gln Ser Leu Arg 1 5 10 15

Leu Glu Tyr Leu Gln Ile Pro Pro Val Ser Arg Ala Tyr Thr Thr Ala 20 25 30

Cys Val Leu Thr Thr Ala Ala Val Gln Leu Glu Leu Ile Thr Pro Phe 35 40 45

Gln Leu Tyr Phe Asn Pro Glu Leu Ile Phe Lys His Phe Gln Ile Trp 50 60

Arg Leu Ile Thr Asn Phe Leu Phe Phe Gly Pro Val Gly Phe Asn Phe 65 70 75 80

Leu Phe Asn Met Ile Phe Leu Tyr Arg Tyr Cys Arg Met Leu Glu Glu 85 90 95

Gly Ser Phe Arg Gly Arg Thr Ala Asp Phe Val Phe 100 105

<210> 342

<211> 23

<212> PRT

<213> Homo sapiens

<400> 342

Leu Ile Phe Lys His Phe Gln Ile Trp Arg Leu Ile Thr Asn Phe Leu

1 10 15

Phe Phe Gly Pro Val Gly Phe

<210> 343

<211> 25

<212> PRT

<213> Homo sapiens

<400> 343

Phe Leu Tyr Arg Tyr Cys Arg Met Leu Glu Glu Gly Ser Phe Arg Gly 1 5 10 15

```
Arg Thr Ala Asp Phe Val Phe Met Phe
             20
<210> 344
<211> 23
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (19)
<223> Xaa equals any of the naturally occurring L-amino acids
Leu Val Phe Leu Gly Gln Ala Phe Thr Ile Met Leu Val Tyr Val Trp
                                     10
Ser Arg Xaa Asn Pro Tyr Val
            20
<210> 345
<211> 21
<212> PRT
<213> Homo sapiens
<400> 345
Val Leu Met Gly Phe Ser Leu Leu Gly Asn Ser Ile Ile Val Asp
                 5
Leu Leu Gly Ile Ala
<210> 346
<211> 25
<212> PRT
<213> Homo sapiens
<400> 346
Asn Gln Pro Gly Gly Ile Arg Ile Leu Lys Thr Pro Ser Ile Leu Lys
Ala Ile Phe Asp Thr Pro Asp Glu Asp
           20
<210> 347
<211> 28
<212> PRT
<213> Homo sapiens
<400> 347
```

PCT/US99/05804 WO 99/47540

173

Arg Leu Glu Tyr Leu Gln Ile Pro Pro Val Ser Arg Ala Tyr Thr Thr

Ala Cys Val Leu Thr Thr Ala Ala Val Gln Leu Glu

<210> 348

<211> 31

<212> PRT

<213> Homo sapiens

<400> 348

Arg Leu Ile Thr Asn Phe Leu Phe Phe Gly Pro Val Gly Phe Asn Phe 1 5 10 15

Leu Phe Asn Met Ile Phe Leu Tyr Arg Tyr Cys Arg Met Leu Glu 25

<210> 349

<211> 12

<212> PRT

<213> Homo sapiens

<400> 349

His Ala Ser Ala Gly Pro Asp Gly Ser Ser Pro Ala 5

<210> 350

<211> 115

<212> PRT

<213> Homo sapiens

<400> 350

Glu Leu Leu Glu Lys Pro Lys Pro Trp Gln Pro Pro Ala Ala Ala

Pro His Arg Ala Leu Leu Val Leu Cys Tyr Ser Ile Val Glu Asn Thr 25

Cys Ile Ile Thr Pro Thr Ala Lys Ala Trp Lys Tyr Met Glu Glu

Ile Leu Gly Phe Gly Lys Ser Val Cys Asp Ser Leu Gly Arg Arg His

Met Ser Thr Cys Ala Leu Cys Asp Phe Cys Ser Leu Lys Leu Glu Gln 70

Cys His Ser Glu Ala Ser Leu Gln Arg Gln Gln Cys Asp Thr Ser His 90

Lys Thr Pro Phe Ala Ala Pro Cys Leu Pro Pro Arg Ala Cys Pro Ser

100 105 110

Ala Thr Arg 115

<210> 351

<211> 77

<212> PRT

<213> Homo sapiens

<400> 351

Leu Pro Gly Trp Gly Phe Pro Thr Lys Ile Cys Asp Thr Asp Tyr Ile
1 5 10 15

Gln Tyr Pro Asn Tyr Cys Ser Phe Lys Ser Gln Gln Cys Leu Met Arg 20 25 30

Asn Arg Asn Arg Lys Val Ser Arg Met Arg Cys Leu Gln Asn Glu Thr 35 40 45

Tyr Ser Ala Leu Ser Pro Gly Lys Ser Glu Asp Val Val Leu Arg Trp 50 55 60

Ser Gln Glu Phe Ser Thr Leu Thr Leu Gly Gln Phe Gly 65 70 75

<210> 352

<211> 65

<212> PRT

<213> Homo sapiens

<400> 352

Ser Pro Val Leu Leu Pro Ala Phe Pro Pro Leu Pro Val Pro Leu Leu 1 5 10 15

Ala Leu Pro Val Ser Ala Pro Leu Pro Ala Cys Val Leu Val Ser Ala
20 25 30

Pro Ala Cys Ala Pro Leu Leu Ala Pro Ala Cys Ala Leu Ala Leu Ala 35 40 45

Pro Gly Phe Pro Gly Thr Arg Arg Ile Val Gly Ala Leu Pro Arg Cys 50 60

Суs

65

<210> 353

<211> 35

<212> PRT

WO 99/47540 PCT/US99/05804

```
<400> 353
Leu Leu Val Leu Cys Tyr Ser Ile Val Glu Asn Thr Cys Ile Ile Thr
Pro Thr Ala Lys Ala Trp Lys Tyr Met Glu Glu Glu Ile Leu Gly Phe
                                 25
             20
Gly Lys Ser
<210> 354
<211> 26
<212> PRT
<213> Homo sapiens
<400> 354
Leu Lys Leu Glu Gln Cys His Ser Glu Ala Ser Leu Gln Arg Gln Gln
Cys Asp Thr Ser His Lys Thr Pro Phe Ala
            20
<210> 355
<211> 40
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (27)
<223> Xaa equals any of the naturally occurring L-amino acids
Gln Val Ser Gly Leu Ile Leu Ser Leu Ser Cys Gly Met Asp Gly Leu
                                    10
Ala Leu Asp Gly Ser Pro Ser Pro Ser Pro Xaa Thr Glu Lys Ala Gly
Arg Cys Ile Ser Gln Thr Ser Leu
<210> 356
<211> 46
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (27)
<223> Xaa equals any of the naturally occurring L-amino acids
```

<400> 356 Gln Val Ser Gly Leu Ile Leu Ser Leu Ser Cys Gly Met Asp Gly Leu

Ala Leu Asp Gly Ser Pro Ser Pro Ser Pro Xaa Thr Glu Lys Ala Gly

Arg Cys Ile Ser Gln Thr Ser Leu Pro Gly Lys Trp Glu Val 40

<210> 357

<211> 173

<212> PRT 1

<213> Homo sapiens

<220>

<221> SITE

<222> (118)

<223> Xaa equals any of the naturally occurring L-amino acids

Arg Ala Ser Lys Thr Val Pro Arg Met Pro Pro Asn Trp Pro Ala Lys

Met Pro Cys Leu Cys His Ile Arg Thr Val Glu His Leu Gly Thr Ile

Ser Ser Gly Ala Pro Gly Arg Pro Thr Gly Gln Gln Ala Ala Arg Thr

Tyr His Ile Cys Trp Ile His Pro Gly Gln Lys Ile Asp Ser Leu Pro

Pro Ser Ser Gln His Pro Arg Ser Gln Gln Leu Ala Pro Gly Thr Trp 70

Pro Ser Thr Ser Thr Thr Lys Pro Ala Glu Glu Thr Leu Gly Ser Ser 90

Ala Ser Leu Pro Ile Ser Gln Ala Arg Lys Ser Glu Lys Cys Thr Phe 100 105

Gln Pro Ser Pro Trp Xaa Val Arg Gly Lys Glu Ser His Gln Val Pro 120

Ala His Pro Ser His Arg Thr Glu Thr Glu Ser Asp His Ser Pro Val 135 130

Arg Lys Pro Pro Ser Arg Gly Thr Arg Thr Gly Asp Phe Thr Val Gly 150 155

Asp Trp Ser Glu Ala Trp Leu Leu Glu Leu Ala Leu Leu 165

```
<210> 358
<211> 23
<212> PRT
<213> Homo sapiens
<400> 358
Arg Met Pro Pro Asn Trp Pro Ala Lys Met Pro Cys Leu Cys His Ile
                      10
Arg Thr Val Glu His Leu Gly
  20
<210> 359
<211> 25
<212> PRT
<213> Homo sapiens
<400> 359
Gly Arg Pro Thr Gly Gln Gln Ala Ala Arg Thr Tyr His Ile Cys Trp
Ile His Pro Gly Gln Lys Ile Asp Ser
    20
<210> 360
<211> 25
<212> PRT
<213> Homo sapiens
<400> 360
Trp Pro Ser Thr Ser Thr Thr Lys Pro Ala Glu Glu Thr Leu Gly Ser
Ser Ala Ser Leu Pro Ile Ser Gln Ala
           20
<210> 361
<211> 23
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (13)
<223> Xaa equals any of the naturally occurring L-amino acids
Lys Ser Glu Lys Cys Thr Phe Gln Pro Ser Pro Trp Xaa Val Arg Gly
                                   10
 1
Lys Glu Ser His Gln Val Pro
```

```
<210> 362
<211> 24
<212> PRT
<213> Homo sapiens
<400> 362
Lys Pro Pro Ser Arg Gly Thr Arg Thr Gly Asp Phe Thr Val Gly Asp
                  5
                                     10
Trp Ser Glu Ala Trp Leu Leu Glu
             20
<210> 363
<211> 10
<212> PRT
<213> Homo sapiens
<400> 363
Pro Cys Ala Asp Cys Leu Ser Ala Trp Ala.
<210> 364
<211> 11
<212> PRT
<213> Homo sapiens
<400> 364
His Ala Ser Gly Tyr Leu Cys Ile Val Leu Leu
<210> 365
<211> 34
<212> PRT
<213> Homo sapiens
<400> 365
Asn Ser Ala Arg Ala Arg Ala Glu Ile Val Leu Gly Leu Leu Val
Trp Thr Leu Ile Ala Gly Thr Glu Tyr Phe Arg Val Pro Ala Phe Gly
Trp Val
<210> 366
<211> 22
<212> PRT
```

```
<213> Homo sapiens
```

<400> 366

Pro Cys Ser Pro Pro Asp Ser Pro Pro Leu Pro Gly Ala Phe Val Trp

1 5 10 15

Arg Val Leu Trp Val Cys
20

<210> 367

<211> 25

<212> PRT

<213> Homo sapiens

<400> 367

Ala Arg Ala Cys Phe Ala Tyr Asn Gly Val Cys Ser Glu Gly Arg Cys

1 10 15

Trp Asp Ser His Phe His Gly Ser Val 20 25

<210> 368

<211> 100

<212> PRT

<213> Homo sapiens

<400> 368

Met Ser Asn Met Gly Lys Ile Pro Ser Leu Ser Leu His Ile Pro Ile
1 5 10 15

Asn Lys Tyr Ile Cys Ser Arg Ile Pro Lys Phe Ile Gln Lys Val Asn 20 25 30

Lys Ser Thr Val Leu Gln Ile Cys Leu Lys Arg Gln Ile Ile Leu Asn $35 \hspace{1.5cm} 40 \hspace{1.5cm} 45$

Lys Asn Lys Met Ser Asp His Ser Lys Ile Gly Lys Ala Asn Leu Val 50 55 60

Gln Ile Asp Ile His Ser Leu Gly Ile Val Glu Thr Gly Cys Val Pro 65 70 75 80

Ser Lys Arg Tyr Cys Thr Leu Leu Thr Glu Gln Ser Gly Phe Pro Phe 85 90 95

Leu Ser His Pro 100

<210> 369

<211> 84

<212> PRT

<213> Homo sapiens

WO 99/47540 PCT/US99/05804

180

<220> <221> SITE <222> (54) <223> Xaa equals any of the naturally occurring L-amino acids <220> <221> SITE <222> (58) <223> Xaa equals any of the naturally occurring L-amino acids <220> <221> SITE <222> (82) <223> Xaa equals any of the naturally occurring L-amino acids <400> 369 Met Ala Gly Cys Cys Leu Lys Leu Phe Gly Val Leu Ser Leu Cys Phe 10 Leu Cys Gly Leu Ile Ser Ile Glu Arg Val Ile Cys Asn Pro Val Ser Ala Asp Phe Gln Val Ser Thr Phe Cys Gln Arg His Cys Leu Leu Arg Ser Lys Val Met Phe Xaa Ile Lys Gly Xaa Thr Ala Thr Ile Glu Val Ile Asn Glu Asn Cys Thr Leu Val Ala Ala Pro Pro Ile Gly Phe Pro Ile Xaa Phe Leu <210> 370 <211> 49 <212> PRT <213> Homo sapiens <400> 370 Met Ser Asp His Ser Lys Ile Gly Lys Ala Asn Leu Val Gln Ile Asp Ile His Ser Leu Gly Ile Val Glu Thr Gly Cys Val Pro Ser Lys Arg

Tyr Cys Thr Leu Leu Thr Glu Gln Ser Gly Phe Pro Phe Leu Ser His

Pro

WO 99/47540

181

```
<210> 371
<211> 50
<212> PRT
<213> Homo sapiens
Met Ala Gly Cys Cys Leu Lys Leu Phe Gly Val Leu Ser Leu Cys Phe
                  5
                                    10
Leu Cys Gly Leu Ile Ser Ile Glu Arg Val Ile Cys Asn Pro Val Ser
Ala Asp Phe Gln Val Ser Thr Phe Cys Gln Arg His Cys Leu Leu Arg
Ser Lys
   50
<210> 372
<211> 34
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (4)
<223> Xaa equals any of the naturally occurring L-amino acids
<220>
<221> SITE
<223> Xaa equals any of the naturally occurring L-amino acids
<221> SITE
<222> (32)
<223> Xaa equals any of the naturally occurring L-amino acids
<400> 372
Val Met Phe Xaa Ile Lys Gly Xaa Thr Ala Thr Ile Glu Val Ile Asn
Glu Asn Cys Thr Leu Val Ala Ala Pro Pro Ile Gly Phe Pro Ile Xaa
                                 25
Phe Leu
```

<210> 373 <211> 65 <212> PRT <213> Homo sapiens

Pro Thr Glu Gly Arg Gln Lys Val Leu Lys Thr Phe Thr Val Pro Arg 10

Ser Ala Leu Ala Met Thr Lys Thr Ser Thr Cys Ile Tyr His Phe Leu

Val Leu Ser Trp Tyr Thr Phe Leu Asn Tyr Tyr Ile Ser Gln Glu Gly 40

Lys Asp Glu Val Lys Pro Lys Ile Leu Ala Asn Gly Ala Arg Trp Lys

Tyr 65

<210> 374

<211> 35

<212> PRT

<213> Homo sapiens

<400> 374

Pro Arg Ser Ala Leu Ala Met Thr Lys Thr Ser Thr Cys Ile Tyr His

Phe Leu Val Leu Ser Trp Tyr Thr Phe Leu Asn Tyr Tyr Ile Ser Gln

Glu Gly Lys

<210> 375

<211> 24

<212> PRT

<213> Homo sapiens

<400> 375

Pro Thr Glu Gly Arg Gln Lys Val Leu Lys Thr Phe Thr Val Pro Arg

Ser Ala Leu Ala Met Thr Lys Thr 20

<210> 376

<211> 27

<212> PRT

<213> Homo sapiens

Phe Leu Asn Tyr Tyr Ile Ser Gln Glu Gly Lys Asp Glu Val Lys Pro

PCT/US99/05804 WO 99/47540

183

```
Lys Ile Leu Ala Asn Gly Ala Arg Trp Lys Tyr
             20
<210> 377
<211> 13
<212> PRT
<213> Homo sapiens
<400> 377
Phe Lys Asp Gln Leu Val Tyr Pro Leu Leu Ala Phe Thr
                 5
<210> 378
<211> 13
<212> PRT
<213> Homo sapiens
<400> 378
Arg Gln Ala Leu Asn Leu Pro Asp Val Phe Gly Leu Val
<210> 379
<211> 10
<212> PRT
<213> Homo sapiens
<400> 379
Ala Thr Ala Ser His Asp Leu Leu Phe
<210> 380
<211> 97
<212> PRT
<213> Homo sapiens
<220>
<221> SITE
<222> (72)
<223> Xaa equals any of the naturally occurring L-amino acids
Met Ser Ile Asn Ile Cys Leu Met Gln Ser Lys Thr Gln Gly Ser Cys
Gln Tyr Leu Leu Pro His Pro Val Pro Ile Ile Leu Lys Val Ser
             20
                                 25
                                                    30
Thr Val Phe Ser Leu Leu Ser Leu Phe Arg Leu Leu Phe Leu Ser Phe
```

Cys Pro His Pro Lys Lys Cys Ser Tyr Leu Leu Lys Tyr Tyr Gly Pro

WO 99/47540 PCT/US99/05804

184

50 55 60

Leu Glu Gly His Lys Thr Leu Xaa Tyr Leu Arg Thr Asn Leu Gly Val
65 70 75 80

Ile Gln Pro Pro Leu Arg Met Tyr Ala Ala Glu Asp Cys Asn Gly Ile
85 90 95

Gly

<210> 381

<211> 46

<212> PRT

<213> Homo sapiens

<400> 381

Met Ser Ile Asn Ile Cys Leu Met Gln Ser Lys Thr Gln Gly Ser Cys

1 10 15

Gln Tyr Leu Leu Pro His Pro Val Pro Ile Ile Leu Lys Val Ser 20 25 30

Thr Val Phe Ser Leu Leu Ser Leu Phe Arg Leu Leu Phe Leu
35 40 45

<210> 382

<211> 51

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (26)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 382

Ser Phe Cys Pro His Pro Lys Lys Cys Ser Tyr Leu Leu Lys Tyr Tyr

Gly Pro Leu Glu Gly His Lys Thr Leu Xaa Tyr Leu Arg Thr Asn Leu 20 25 30

Gly Val Ile Gln Pro Pro Leu Arg Met Tyr Ala Ala Glu Asp Cys Asn 35 40 45

Gly Ile Gly 50

<210> 383

<211> 23

<212> PRT

```
<213> Homo sapiens
 <400> 383
 Lys Glu Glu Asp Asp Asp Thr Glu Arg Leu Pro Ser Lys Cys Glu Val
 Cys Lys Leu Leu Ser Thr Glu
             20
 <210> 384
 <211> 23
 <212> PRT
 <213> Homo sapiens
 Lys Glu Glu Asp Asp Asp Thr Glu Arg Leu Pro Ser Lys Cys Glu Val
                                    10
 Cys Lys Leu Leu Ser Thr Glu
             20
 <210> 385
 <211> 19
 <212> PRT
 <213> Homo sapiens
 <400> 385
 Leu Gln Ala Glu Leu Ser Arg Thr Gly Arg Ser Arg Glu Val Leu Glu
 Leu Gly Gln
 <210> 386
 <211> 19
 <212> PRT
 <213> Homo sapiens
 <400> 386
 Leu Gln Ala Glu Leu Ser Arg Thr Gly Arg Ser Arg Glu Val Leu Glu
                  5
                                    10
 Leu Gly Gln
 <210> 387
<211> 12
<212> PRT
 <213> Homo sapiens
 <400> 387
```

Arg Gln Ala Val Ile Val Cys Arg Arg Arg Phe Val
1 5 10

<210> 388
<211> 148

<212> PRT <213> Homo sapiens

<400> 388

Pro Pro Arg Trp Ala His Pro Lys Ala Pro Glu Gly Ser Pro Asp Pro 1 5 10 15

Pro Ser Pro Pro Ser Ala Leu Gly Leu Ser Val Leu Pro Trp Ser Asp 20 25 30

Ser Asp Pro Trp His Ile Ser Val Ser Pro Cys Ala Gln Arg Glu His

Tyr Ser Pro Gly Ser Ala His Ile Asn Ser Leu Arg Pro Leu Pro Ala 50 55 60

Leu Ser Leu Lys Arg Cys Lys Ala Arg Val Ser Ser Ser Cys Leu Tyr 65 70 75 80

Pro Ala Pro Ala Pro Ala Pro Leu Glu Ile Asp Arg Cys Asp 85 90 95

Ser Val Pro Pro Val Ala Leu Cys Ser Ala Ala Tyr Thr Leu Arg Ile 100 105 110

Cys Trp Ala Ser Val Leu Cys His Arg Pro Pro Pro Ser Thr Ser Gln 115 120 125

Pro Lys Pro Arg Ala Arg Pro Lys Lys Gly Lys Ala Ile Phe Pro Thr 130 135 140

Ala Gln Val Pro 145

<210> 389

<211> 71

<212> PRT

<213> Homo sapiens

<400> 389

Pro Pro Arg Trp Ala His Pro Lys Ala Pro Glu Gly Ser Pro Asp Pro 1 5 10 15

Pro Ser Pro Pro Ser Ala Leu Gly Leu Ser Val Leu Pro Trp Ser Asp 20 25 30

Ser Asp Pro Trp His Ile Ser Val Ser Pro Cys Ala Gln Arg Glu His
35 40 45

Tyr Ser Pro Gly Ser Ala His Ile Asn Ser Leu Arg Pro Leu Pro Ala 55

Leu Ser Leu Lys Arg Cys Lys

<210> 390

<211> 77

<212> PRT

<213> Homo sapiens

<400> 390

Ala Arg Val Ser Ser Ser Cys Leu Tyr Pro Ala Pro Ala Pro Ala Pro

Ala Pro Leu Glu Ile Asp Arg Cys Asp Ser Val Pro Pro Val Ala Leu 25

Cys Ser Ala Ala Tyr Thr Leu Arg Ile Cys Trp Ala Ser Val Leu Cys 40

His Arg Pro Pro Pro Ser Thr Ser Gln Pro Lys Pro Arg Ala Arg Pro 55

Lys Lys Gly Lys Ala Ile Phe Pro Thr Ala Gln Val Pro 70

<210> 391

<211> 26

<212> PRT

<213> Homo sapiens

<400> 391

Glu Glu Lys Leu Phe Thr Ser Ala Pro Gly Arg Asp Phe Trp Val Met

Gly Glu Thr Arg Asp Gly Asn Glu Glu Asn 20

<210> 392

<211> 42

<212> PRT

<213> Homo sapiens

Gln Lys Pro Thr Phe Ala Leu Gly Glu Leu Tyr Pro Pro Leu Ile Asn

Leu Trp Glu Ala Gly Lys Glu Lys Ser Thr Ser Leu Lys Val Lys Ala 25

PCT/US99/05804

188

Thr Val Ile Gly Leu Pro Thr Asn Met Ser 35 40

INTERNATIONAL SEARCH REPORT

International application No. PCT/US99/05804

A. CLASSIFICATION OF SUBJECT MATTER			
IPC(6) :Picase See Extra Sheet. US CL :Picase See Extra Sheet.			
According to International Patent Classification (IPC) or to both national classification and IPC			
B. FIELDS SEARCHED			
Minimum documentation searched (classification system followed by classification symbols)			
U.S. : 536/23.5; 435/320.1, 252.3, 69.1, 6, 7.1; 530/300, 388.22; 514/2; 436/501			
Documentation searched other than minimum documentation to the exteat that such documents are included in the fields searched			
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)			
Please See Extra Sheet.			
C. DOCUMENTS CONSIDERED TO BE RELEVANT			
Category*	Citation of document, with indication, where ap	opropriate, of the relevant passages	Relevant to claim No.
X	Database EST: Accession No.:AA331279. ADAMS et al. Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence. Nature 377 (6547 Suppl), 3-174 (1995). Nucleotides 1-314.		1-12, 14-21
A	SCHWARTING, R. et al. Bioch purification of human B Cell Stimula Immunol. 1985, Vol. 15, pages 632-6	atory Factor (BSF). Eur. J.	1-21
Further documents are listed in the continuation of Box C. See patent family annex.			
"Special estegories of cited documents: "T" letter document published after the international filing data or priority data and not in conflict with the application but cited to understand			
	cument defining the general state of the art which is not considered be of particular relevance	the principle or theory underlying the	invention
"E" earlier document published on or after the international filing data "X" document of particular relevance; considered novel or cannot be come:			
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other		when the document is taken alone	
special reason (as specified) *O* document referring to an oral disclosure, use, exhibition or other means		"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art	
*P" document published prior to the international filing date but later than *& the priority date claimed		·	
Date of the	actual completion of the international search	Date of mailing of the international sea	rch report
07 JUNE 1999		07 JUL 1999	
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks		Authorized officer Laurence	1
Box PCT Washington, D.C. 20231		EILEEN O'HARA	you
• •		Telephone No. (703) 308-0196	

INTERNATIONAL SEARCH REPORT

International application No. PCT/US99/05804

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)			
This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:			
1. Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely:			
Claims Nos.: because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:			
3. Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).			
Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)			
This International Searching Authority found multiple inventions in this international application, as follows:			
Please See Extra Shoot.			
As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.			
2. As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.			
As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:			
4. X No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: 1-21			
Remark on Protest			
No protest accompanied the payment of additional search fees.			

INTERNATIONAL SEARCH REPORT

International application No. PCT/US99/05804

A. CLASSIFICATION OF SUBJECT MATTER:

C07H 21/04; C07K 1/00, 16/00; C12N 15/00, 1/20; C12P 21/06; A61K 38/00; C12Q 1/68; G01N 33/53, 33/566

A. CLASSIFICATION OF SUBJECT MATTER: US CL :

536/23.5; 435/320.1, 252.3, 69.1, 6, 7.1; 530/300, 388.22; 514/2; 436/501

B. FIELDS SEARCHED

Electronic data bases consulted (Name of data base and where practicable terms used):

SEQUENCE DATA BASE MPSRCH: EST, GenEmbl, N_Geneseq_34, Issued_Patents_NA, SPTREMBL_8, SwissProt_36, PIR_58_Issued Patents_AA, (SEQ ID NOS: 11 and 108 only). One nucleotide sequence and one amino acid sequence have been searched. It is not clear which sequences are embraced by the claims because the claims refer to sequences X and Y. The table at pages 180-188 contains many sequences X and Y, yet the claims refer to X and Y in the singular only. Accordingly, the first X nucleotide sequence disclosed and the first Y amino acid sequence disclosed were searched.

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claims 1-21, drawn to nucleic acid molecules, vectors, host cells containing recombinant nucleic acid molecules, polypeptides, antibodies, a method of producing a polypeptide, a method for treating a medical condition comprising administering a polypeptide, a method of diagnosing a pathological condition by genetic analysis or protein assay, and a method for identifying a binding partner to a polypeptide, for gene 1, the nucleic acid molecule identified by SEQ ID NO:11, and the polypeptide identified by SEQ ID NO: 108, as listed in the table on pages 180-188 of the description. Groups II through XCV, claims 1-21 for each group, drawn to nucleic acid molecules, vectors, host cells containing recombinant nucleic acid molecules, polypeptides, antibodies, a method of producing a polypeptide, a method for treating a medical condition comprising administering a polypeptide, a method of diagnosing a pathological condition by genetic analysis or protein assay, and a method for identifying a binding partner to a polypeptide for, genes 2 through 95, the nucleic acid molecules identified by SEQ ID NOS:12 through 105, and the polypeptides identified by SEQ ID NOS:109 through 202 respectively, as listed in the table on pages 180-188 of the description.

The inventions listed as Groups I through XCV do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: Pursuant to 37 C.F.R.§ 1.475(b-d), the ISA/US considers that where multiple products and processes are claimed, the main invention shall consist of the first invention of the category first mentioned in the claims and the first recited invention of each of the other categories related thereto.

Pursuant to 37 C.F.R. § 1.475(b-d), the ISA/US considers that where multiple products and processes are claimed, the main invention shall consist of the first invention of the category first mentioned in the claims and the first recited invention of each of the other categories related thereto. Accordingly, the main invention (Group I) comprises the first recited product, a polynucleotide comprising gene No. 1, identified by SEQ ID NO:11, the polypeptide it encodes, identified by SEQ ID NO:108, methods of producing the polypeptide, and methods of using the polynucleotide and polypeptide. Further pursuant to 37 C.F.R. § 1.475(b-d), the ISA/US considers that any feature which the subsequently recited products and methods share with the main invention does not consitute a special technical feature within the meaning of PCT Rule 13.2 and that each of such products and methods accordingly defines a separate invention.

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

BLACK BORDERS

IMAGE CUT OFF AT TOP, BOTTOM OR SIDES

FADED TEXT OR DRAWING

BLURRED OR ILLEGIBLE TEXT OR DRAWING

SKEWED/SLANTED IMAGES

COLOR OR BLACK AND WHITE PHOTOGRAPHS

GRAY SCALE DOCUMENTS

LINES OR MARKS ON ORIGINAL DOCUMENT

REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY

IMAGES ARE BEST AVAILABLE COPY.

☐ OTHER: _____

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.